

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:           October 26, 2001, 18:41:07 ; Search time 6793.37 Seconds  
                  (without alignments)  
                  11987.820 Million cell updates/sec

Title:            US-09-454-684A-174  
Perfect score:    5265  
Sequence:         I gcaatcatgaaatgctgtc.....gtgctcgatgatcttctaa 5265

Scoring table:    IDENTITY\_NUC  
                  Gapop 10.0 , Gapext 1.0

Searched:         1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters:       2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
                  Maximum Match 100%  
                  Listing first 45 summaries

Database :        GenEmbl:\*  
                  1: gb\_bal:\*  
                  2: gb\_ba2:\*  
                  3: gb\_ba3:\*  
                  4: gb\_in1:\*  
                  5: gb\_in2:\*  
                  6: gb\_in3:\*  
                  7: gb\_om:\*  
                  8: gb\_ov:\*  
                  9: gb\_pat1:\*  
                  10: gb\_pat2:\*  
                  11: gb\_ph:\*  
                  12: gb\_pl1:\*  
                  13: gb\_pl2:\*  
                  14: gb\_pl3:\*  
                  15: gb\_pl4:\*  
                  16: em\_bal:\*  
                  17: em\_ba2:\*  
                  18: em\_fun:\*  
                  19: em\_higo\_hum:\*  
                  20: em\_higo\_inv:\*  
                  21: em\_higo\_rod:\*  
                  22: em\_hig\_hum1:\*  
                  23: em\_hig\_hum2:\*  
                  24: em\_hig\_hum3:\*  
                  25: em\_hig\_hum4:\*  
                  26: em\_hig\_hum5:\*  
                  27: em\_hig\_hum6:\*  
                  28: em\_hig\_hum7:\*  
                  29: em\_hig\_hum8:\*  
                  30: em\_hig\_inv1:\*  
                  31: em\_hig\_inv2:\*  
                  32: em\_hig\_other:\*  
                  33: em\_hig\_rod:\*  
                  34: em\_hum1:\*  
                  35: em\_hum2:\*  
                  36: em\_hum3:\*  
                  37: em\_hum4:\*  
                  38: em\_hum5:\*  
                  39: em\_hum6:\*  
                  40: em\_hum7:\*  
                  41: em\_in:\*  
                  42: em\_om:\*  
                  43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v1:\*  
59: gb\_v12:\*  
60: gb\_hgt1:\*  
61: gb\_hgt2:\*  
62: gb\_hgt3:\*  
63: gb\_hgt4:\*  
64: gb\_hgt5:\*  
65: gb\_hgt6:\*  
66: gb\_hgt7:\*  
67: gb\_hgt8:\*  
68: gb\_hgt9:\*  
69: gb\_hgt10:\*  
70: gb\_hgt11:\*  
71: gb\_hgt12:\*  
72: gb\_hgt13:\*  
73: gb\_hgt14:\*  
74: gb\_hgt15:\*  
75: gb\_hgt16:\*  
76: gb\_hgt17:\*  
77: gb\_hgt18:\*  
78: gb\_hgt19:\*  
79: gb\_hgt20:\*  
80: gb\_hgt21:\*  
81: gb\_hgt22:\*  
82: gb\_hgt23:\*  
83: gb\_hgt24:\*  
84: gb\_hgt25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_rol:\*  
95: gb\_rod:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5180.8	98.4	14973	1	AE001314
2	2456.6	46.7	12838	1	AE002338
3	824.2	15.7	10199	1	AE001315
4	435.8	8.3	5172	16	AB033817
5	435.8	8.3	14781	1	AE001638
6	435.8	8.3	20517	1	AE002181
7	435.8	8.3	30050	2	AP002547
8	61.8	1.2	117655	75	AC074221

9	61	1.2	1159	89	AF126748	Homo sapi	AF126748 Homo sapi
c 10	61	1.2	1442	88	AF087653	Homo sapi	AF087653 Homo sapi
c 11	61	1.2	1472	88	AF126749	Homo sapi	AF126749 Homo sapi
12	60.4	1.1	720	85	AB007820	Homo sapi	AB007820 Homo sapi
13	60.4	1.1	1410	85	AB009843	Homo sapi	AB009843 Homo sapi
14	60.4	1.1	2479	95	RN087960	Rattus norv	U87960 Rattus norv
c 15	58.8	1.1	624	5	AF139019	Cepaea ne	AF139019 Cepaea ne
16	57.6	1.1	222172	75	AC079245	Mus muscu	AC079245 Mus muscu
17	57.6	1.1	471	97	HSU23863	Human clone	U23863 Human clone
c 18	57.6	1.1	1133	53	CNS062LW	Human T3 end of	AL422490 T3 end of
19	57.6	1.1	3103	97	HSU80742	Homo sapien	U80742 Homo sapien
20	57.6	1.1	4751	91	BC004354	Homo sapi	BC004354 Homo sapi
21	57.6	1.1	5120	10	AX067148	Sequence	AX067148 Sequence
22	57.6	1.1	5120	89	AF152102	Homo sapi	AF152102 Homo sapi
23	57.6	1.1	6604	91	D83783	Human mRNA	D83783 Human mRNA
24	57.6	1.1	6757	89	AF117755	Homo sapi	AF117755 Homo sapi
25	57.6	1.1	6788	88	AF071309	Homo sapi	AF071309 Homo sapi
c 26	57.4	1.1	186838	67	AC032386	Mus muscu	AC032386 Mus muscu
27	57.2	1.1	195	9	AR002263	Sequence	AR002263 Sequence
28	57.2	1.1	195	9	AR053097	Sequence	AR053097 Sequence
29	57.2	1.1	234	9	AR002264	Sequence	AR002264 Sequence
30	57.2	1.1	234	9	AR053098	Sequence	AR053098 Sequence
31	57	1.1	299	92	HSAL31896	Homo sapi	AJ131896 Homo sapi
32	56.8	1.1	323	8	SSA012206	Salmo sal	AJ012206 Salmo sal
33	56.6	1.1	154	9	AR002267	Sequence	AR002267 Sequence
34	56.6	1.1	154	9	AR053101	Sequence	AR053101 Sequence
35	56.6	1.1	168	9	AR002265	Sequence	AR002265 Sequence
36	56.6	1.1	168	9	AR053099	Sequence	AR053099 Sequence
37	56.6	1.1	171	9	AR002266	Sequence	AR002266 Sequence
38	56.6	1.1	171	9	AR053100	Sequence	AR053100 Sequence
39	56.6	1.1	477	10	E34519	SCA7 gene a	E34519 SCA7 gene a
40	56.6	1.1	477	88	AF020275	Homo sapi	AF020275 Homo sapi
41	56.6	1.1	197897	82	AP000779	Homo sapi	AP000779 Homo sapi
42	56.6	1.1	215963	71	AC040982	Mus muscu	AC040982 Mus muscu
c 43	56.2	1.1	326419	76	AC079635	Sequence	AC079635 Mus muscu
44	56	1.1	172	97	S82497	SCA1-spinoc	S82497 SCA1-spinoc
45	56	1.1	471	93	HSCAG40	H. sapiens m	X85325 H. sapiens m

ALIGNMENTS

RESULT	1
AE001314	
LOCUS	AE001314 14973 bp DNA BCT 30-OCT-2000
DEFINITION	Chlamydia trachomatis section 41 of 87 of the complete genome.
ACCESSION	AE001314 AE001273
VERSION	AE001314.1 GI:3328833
KEYWORDS	
SOURCE	Chlamydia trachomatis.
ORGANISM	Chlamydia trachomatis.
REFERENCE	1 (bases 1 to 14973)
AUTHORS	Stephens, R.S., Kalman, S., Lammel, C.J., Fan, J., Marathe, R., Aravind, L., Mitchell, W.P., Olinger, L., Tatusov, R.L., Zhao, Q., Koonin, E.V. and Davis, R.W.
TITLE	Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
JOURNAL	Science 282 (5389), 754-759 (1998)
MEDLINE	99000809
PUBMED	9784136
REFERENCE	2 (bases 1 to 14973)
AUTHORS	Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W., Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
TITLE	Comparative genomes of Chlamydia pneumoniae and C. trachomatis
JOURNAL	Nat. Genet. 21 (4), 385-389 (1999)
MEDLINE	99206606
PUBMED	10192388
REFERENCE	3 (bases 1 to 14973)
AUTHORS	Stephens, R.S., Kalman, S., Lammel, C.J., Fan, J., Marathe, R., Aravind, L., Mitchell, W.P., Olinger, L., Tatusov, R.L., Zhao, Q., Koonin, E.V. and Davis, R.W.
TITLE	Direct Submission

JOURNAL	Submitted (20-MAY-1998) Program in Infectious Diseases, University of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA
FEATURES	
source	Location/Qualifiers
gene	1..14973
CDS	/organism="Chlamydia trachomatis"
	/strain="D/UW-3/CX"
	/db_xref="taxon:813"
	228..692
	/gene="CT406"
	228..692
	/gene="CT406"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAC68003.1"
	/db_xref="GI:3328834"
	/translation="MLPCFNHGELKVIDSRNAPESNAIKRRRCLRCRSORFTTFETV
	ELTVQVLKRGDRVENFQESKLVAASHSHTRIGQEQVAIASNIKODLLGKONREI
	STKEIGELVMKYLKKADMIAYIRFACVYRRFKDVGELMEVLLSATPDGEQT"
	706..1080
	/gene="dksa"
	/note="CT407"
	706..1080
	/gene="dksa"
	/codon_start=1
	/transl_table=11
	/product="DnaK Suppressor"
	/protein_id="AAC68004.1"
	/db_xref="GI:3328835"
	/translation="MPLTDEEIANFKTLLEMAKAKLSHTLEGNAQEVKKPNEATGYSO
	HPAQDQDTDFDTISLEVTWKYKLRQIDRALEKIEEASYGICDVSGETPLARLMA
	IYATMTVKSQEKEFKGLSGN"
	1086..1589
	/gene="lspa"
	/note="CT408"
	1086..1589
	/gene="lspa"
	/codon_start=1
	/transl_table=11
	/product="Lipoprotein Signal Peptidase"
	/protein_id="AAC68005.1"
	/db_xref="GI:3328836"
	/translation="MPNRSPLFTLLLLASIDWVKLVLLKSCQLSPHSSAFLYSY
	VWGHFSLIIPSEFNEGAAGLFTQYKIPLIIFRVCVILGLALFRIKYSLSHRRTRVA
	LTLTLAGLGNVGDIILYGVDFLSLSYYSWRPFSENLADAFISIGTLLLIHLHYFT
	KESKYP"
	1691..3052
	/gene="CT409"
	1691..3052
	/gene="CT409"
	/codon_start=1
	/transl_table=11
	/product="Amino acid permease"
	/protein_id="AAC68006.1"
	/db_xref="GI:3328837"
	/translation="MSAFSLPFLAFDDFIWSYIAFVILLGLVLTCKSKFAQFTQLP
	AFKFLFYHFSKEASKKNGKGVPLKVPFASAGNIGNVVGIIVTAASVGGPGAL
	FWWVAGLSIVYKSEVYLGKFRQSDANNVHGGPMFFLDKAYRIRIVSVIYAVLL
	CIGVEIYQFISITLASCWDVPLMAISLLFLVMFVAVGGQGRICKICSLVLPFF
	MLVYGMAFYILIOEIHITLPALFSIFRSFTGHTAGTGGAGCTVASTIRGLSRAAY
	SGDIGIGDSIIQSETATNPQAOISIVGIVDNLVCTLSLLIVLASGVWHRAGLE
	SGEIVEQALSFPYIRVFLPAFLFATGYTTIISFLVGVKKCANFICGNKSKFTYV
	GIIALPAFCFLPQDTALLVMSVGALLLCLNLLGVFLMRKELIFPKKEAETLEVSEST
	LSS"
	3267..4544
	/gene="pcnB_1"
	/note="CT410"
	3267..4544
	/gene="pcnB_1"
	/codon_start=1
	/transl_table=11
	/product="PolyA Polymerase"
	/protein_id="AAC68007.1"

```

/db_xref="GI:3328838"
/translation="MNAQCNELLSRGLDLDSKNHSHNLTPITFYPSDHNIDLOQSFST
HALSVRTLKAGYAYIVGGCIRDLDLKTEPKDFDITSAKPEVKTLFKNCILVKG
RFLRAHFPNQLIEVATFRSGNEEDSLITKDNLSAGAEEDVLRDDETINGLFDEPLS
AELIIDYGVVEDLKNRYLRTIGDPLRFPKQDPVRLMLKLLILSYATVDPKTLEDS
QESCHELVKSOPRVEEIIKVLSSGESTIFFQLASERHLEILFPYMAKAFSLSKTL
QDQTFACALTDKIKRKSFCSEHFLLAILLFFIVNFNVRKYKSYQYPTMTIQSIFDY
IRNFTEFFADFSCKKNFTLLTLQMQRYLPLNAMKKEKKSFTNRRLLRHAYE
REALYLLIRSVKVPVEATYGEWLKHVELASE"
gene
4605..6428
/genes="lpxB"
/notes="CT411"
4605..6428
/genes="lpxB"
/codon_start=1
/transl_table=11
/product="Lipid A Disaccharide Synthase"
/protein_id="AAC68008.1"
/db_xref="GI:3328839"
/translation="MFPQKITLWLYPLGLFANLFGTAFPCVQMSLTKKGYSVVPKIF
WYLSGTAVGICHGFIQSQYPIALLHFNLIYIPRNLIASLNPSPKIASLSEVSV
ATAITVSAIGTRPLPHMTWASPNIHLNLPESLSWOLIGICIGITPISLRFITQWF
VLEYKNOALPAPFKWASLGGICLLYFLRTGDVNLVLCYCGFLPSSLANLRIASRE
AFRKPSNSCFIACSEHSGDTLGGNLLKEMHAKYPDHICFVGVGQPMRAQNFALFAM
EKQVSGFEWELLAPLKWRYQLLYNLILNTPRTVICIDFPDFHLLIKLRSRGY
KGIYVYVCPISLWNRPRKTVLEKYLDDLLLLIFEQNLFKDSLALRTVYLGHPLEST
IKSFSPNLWMDKLHLPLDKPPIAAGPGRSRSDILRLNETIQVAFQASLSLTHHLLV
SSANPEYDHLILEVQNRCLSHIVPSQFVRLMRECDFAKCGTIVLEFALNLP
TIVTQCLRPDLFLAKYFINILPAYSLPNIILGRTIRPEFVGKKDKQYEDVAAALN
ILKTSQAOEKDKDSRDYVQAINESASSMKCELSLIFETAS"
6535..9462
/genes="pmpA"
/notes="CT412"
6535..9462
/genes="pmpA"
/codon_start=1
/transl_table=11
/product="Putative outer membrane protein A"
/protein_id="AAC68009.1"
/db_xref="GI:3328840"
/translation="MNRVIEHAHYDQRLQSPNTNLFVHPHYLTLPKFLGLGALIV
YAPYSFAMELAIHGKQKQKQDFTFMISSCPEGNYIINRKLISDFSLLNKVSSGG
AFRNJAGLISFVGNKSSASIHFKHININGFAGVFSESSIEFTDLURKLVAFSESTGG
ITAKEDIKFNKNNHIAFRNNITKGNVGIQLOGDMKGSVSVFDQRGAIIEFTNNQAVT
SSSMKHSRGGAISGDFAGSRILFLNQOITPEGNSAVHGGAIYNKNGVLEPGLNAGP
LAFKENTYIANGGAIYTSNFRANQOTSPILFSONHANKGGAIYAQYVNLQONQDTIR
PEKNTAKEGGAIYTSQCSITAHNTIIFSDNAAQDLGGAILLEGCKPSLFLIASHGN
TAFSGNTMLHTTKASLRHNSILIKEAPYKQLAANKNHSIHFFDPVWALSASSPI
QINAPEYETFPFSPKGMIVFSGANLLDDAREDVANRTSIFNQPVHLYNGTUSIENGHA
LIVQSFQKQGRISLPSGSSALATYMNSSFHGNISSEKPELNGLSFGVDSPSNLQA
EIRAGNAPLRLSGSPSIHDPEGLFYENRDTAASPYQMEILLITSDKIVDISKPTTDSLVA
TNKQSGFQGANHFSQNPNTINNTKOKILRASWLPTEYVLESNRVGRAPVSLGWSTFL
LLQTAASHNLGHLCNNSLIPYTSFVLIGLTGAEMSTHSSPEESFISRLGATGYSII
RLTSLTLISGGSHMFGSFDVADLPEHITSEGIQVONVGLTHWGLPTVNSTLCAALDH
NAMWRICKDHTYKQWDTFGMRGTGLGASYTFLEYDDQTMRFVSPANIBATILQRAFT
ETGYNPFBSFKTLNLNIAPIGIGYFCLGNSSFALLKGGISGYSRDKRENPSYLAH
LAMNDFAWTNGCSYPTSAHTLANOLILRYKACSLIYIATYINREKGNLSNLSGCGY
VGF"
9601..14856
/genes="pmpB"
/notes="CT413"
9601..14856
/genes="pmpB"
/codon_start=1
/transl_table=11
/product="Putative outer membrane protein B"
/protein_id="AAC68010.1"
/db_xref="GI:3328841"
/translation="MKWLSATAVFAAVLPVSVGFCFPEPKELNFSRVGTSSSTFTFET
VGEAGAEIVSGNASFTKFTNIPDTDTPTNSNSSNGSETASVSESDSTPTTTPDP
KGGGAFYNAHSGVLSFMTFRSGTEGSLTSEIKITGEGGAIIFSQGBELLTDLTGLIQN
NLSQSGGAIIGESTISLUSGITKATFSSNSAEVPAVPKRPTEPQAOTRATSVSESSSS
GNDVSVPSSSSRAEPAAANLQSHFICATATPAAQPTDTESTPSHPKPGSGGAIYAKGDL

```

```

TIADSQEVLFSINKATKDGAIFAEKDVSVFENITSLKVQTINGAEKGGAIYAKGDLISI
QSSKQSLNSNYSKQGGALYVEGDIQFODLEIRIKYKPAKGTFFETKTLPKAQAASA
GNADAWSSSQSGGATVSVNSGSSSDSDTSETVPKATAGTFFETKTLTKNLSITNI
TGIIEIANKNATDVGGAIVKCTLFCENSHRLQFLKNSDKQGGIYGEDNITLNLIT
GTLQENKATEEGGGLFIKGTDKALITMTGLDSFCLINNTSEKHGSGAFVKEISQTY
TSDVETIFGITPVHGETVITGNKSTGGNGGVCTKRLASLNLSQISISGNSAENGSG
AHTCSDPSTADTAQPAASAATSPESAPVYSTALSTPSSVSSLTLLAASSQAS
PATSNKETQDPNADTLLIDYVDDTISKNTAKGGGIYAKKAKM3RIDOLNISNSA
TPIGGICCKESLELDALVSLVTENLKGEGGLHAKVTNINLKSQFSFSSNNKANS
SSTGVATTASAPAAAASLQAAAAPVSPATPTYGVVGGAIYGEKVTFSQCSCTCQ
FSGNOAIDNPNSSQSLNVQGGAIYAKTSLISIGSSDAGTSYIFSGNSVSTGKSTQTTGQI
AGGAIYSPVTTLNCPATFSNNTATSMATPKTSSSDSGSNIKDTIGGAIATGATLISG
YSRFGNTADLGAATLANATPSATSONSITEKITLENGSIFERNQANKRGAI
YSPSVSIKGNITFNQNTSHDGSAYFTKDATIESLGSVLTGNHVTQASSATSG
QONTANYGAAIFGDPGTQSSODAILTLASSGNITFSSNLSUHOCDTPASKPCS
TAGYVKLSLQAAKAGKTIISFFDCVHTSKKIGSTONVETLDINKENSNPYTGTVES
SELHENKSLYIPONAILHNGTLVLEKTELHVVSFEQKESKLIMKPGAVLSNQNTANG
ALVINGTLIDLSMGTPOQAEIFSPPELRIYVTSASGGSVSSIPNPKRISAAA
PSGSAATTPPMSENKVFUTGDLTLIDPNNGFYONPMLGSDLDVPLIKDPTNDSVOVY
DUTLSGDLFPQKGYMTWLTDSNPQTKLQARWTFDYRWYVIPIDNHFIANSILGS
QMSMTVVKQGLIINMLNARFDIAYNNFWSGVGTFLAQOQTPLSEFYSYSRGTSV
AIDAKPRODFILGAAFSKMVGKTKAIKNMHWFKHGVSEYQVQSVYGGKFLVFLNKO
HGMALPELIOGVVSYCHLKHDTTLYPSTHERNKGDMEDLGLADLRISMDLKEPSKD
SKRITVIGELEYSSIROKQFTEIDYDRHFDQDVCAYRNLSLPVCAVEGAINMNCITLM
YNKALAYMPSIYRNPNVCYRVLSSNEAGVVICGVPRTSARAETSTQLYLGPFWTL

Query Match      98.4%; Score 5180.6; DB 1; Length 14973;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 5226; Conservative 0; Mismatches 27; Indels 9; Gaps 2;

QY 7 atgaataaggctcagctactcggtggtttgtgtgtctctccctcagttcaggtttt 66
|||||
Db 9601 ATGAAATGGCTGTCAGCTACTGCGGTGTTGCTGCTGTTCTCCCTCAGTTTCAGGCTT 9660

QY 67 tgcctccagaacctaagaataattctctcgcgtagaagaactctctctctacact 126
|||||
Db 9661 TGCTCCCGAGAACCTAAAGAAATTAATTTCTCTCGCGTAGGAACCTTCTTCTCTACCACT 9720

QY 127 tttaactgaacaattggagaagctggggcagaaatataatcgctctctggttaacgcatcttc 186
|||||
Db 9721 TTTACTGAACAGTTGAGAGAGCTGGGAGAGATATATATCGTCTCTGTAACGCATCTTTC 9780

QY 187 acaaaattaccacaattcctactaccgatacaacaactcccaacgactcaaaactcctct 246
|||||
Db 9781 ACAAAATTTACCAACATTCCTACTACCGATAACAACAACCTCCACGAACCTCAACTCCTCT 9840

QY 247 agctctagcggagaaactgcttcggtttcttgagatagtagctctcaacaacagactcct 306
|||||
Db 9841 AGCTCTAACGGAGAGACTGCTTCCGTTTCTGAGGATAGTGACTCTACAACAACGACTCCT 9900

QY 307 gatcctaaagggtggcggcgccttttataacgcgcactccggaggttttctctcttatgaca 366
|||||
Db 9901 GATCCTAAAGTGGCGCGGCCITTTTATACGCGCATCTCCGGAGTTTATCTCTTATGACA 9960

QY 367 cgatcaggaaacagaagtttctcttaactctgtctgagataaaaaatgactggtgaagcggt 426
|||||
Db 9961 CGATCAGGAACAAGAAGTTCTTAACTCTGCTGAGATAAAAAATAACTGCTGAAGGCGGT 10020

QY 427 gctattctctcaagagagactgctatttacaagatctgacaagcttaaccatccaaaat 486
|||||
Db 10021 GCTATCTTCTCAAGGAGAGCTGCTATTATTACAGATCTGACAGGCTTAACCAATCCAAAAT 10080

QY 487 aacttatccagctatccggaggagcgatttttggaggagctctacaatctccctcatcaggg 546
|||||
Db 10081 AACTTATCCAGCTATCCGGAGGAGCGATTTTGTGGAGATCTACAAATCTCCCTATCA3GG 10140

QY 547 attactaaagcgactttctcctcgcaactctgcagaagtttctctgctcctctgttaagaaacct 606
|||||
Db 10141 ATTACTAAAGCGACTTTCTCCTCCAACCTCTGCAGAAGTTCTCTGCTCTGTTAAGAAACT 10200

QY 607 acagaaacctaaagctcaaacagcaagcaagtcgggttcttagtagtcttagcggaat 666
|||||
Db 10201 ACAGAAACCTAAAGCTCAAAACAGCAAGCAACGTCGGGTCTTAGTAGTCTTAGCGGAAAT 10260

```

Qy	667	gattcgggtgtcttcccccagttccagtagagctgaaccccgagcagctaa	tcttcaaagt	726
Db	10261	gattcgggtgtcttcccccaggttccagtagagctgaaccccgagcagctaa	tcttcaaagt	10320
Qy	727	cacittatttgtctacagctactcctgtctcctaaccgcgatacagaatac	taactccc	786
Db	10321	cactttatttgtctacagctactcctgtctcctaaccgcgatacagaatac	taactccc	10380
Qy	787	tctcataagccagagatctggggagctatctatgctaaagcgacctactat	cgcagac	846
Db	10381	tctcataagccagagatctggggagctatctatgctaaagcgacctactat	cgcagac	10440
Qy	847	tctcaagaggtactattctcctaataaataaagctactaaagatggagcgc	gatctttgct	906
Db	10441	tctcaagaggtactattctcctaataaataaagctactaaagatggagcgc	gatctttgct	10500
Qy	907	gagaaaagatgtttottttgcagaaattacatacaataaaaagtcataac	taacgggtgc	966
Db	10501	gagaaaagatgtttottttgcagaaattacatacaataaaaagtcataac	taacgggtgc	10560
Qy	967	gaaaaggagagctatctatgctaaaagtgacctctcaattcaatctctaa	acaacagt	1026
Db	10561	gaaaaggagagagctatctatgctaaaagtgacctctcaattcaatctctaa	acaacagt	10620
Qy	1027	ctttttaattctaactacagtaaaacaagtggtggggcctctatatgttga	aggaggtata	1086
Db	10621	ctttttaattctaactacagtaaaacaagtggtggggcctctatatgttga	aggaggtata	10680
Qy	1087	aacttccaagatcttgaagaattcgcaataaagtcataaagcttgaaacaa	ca	1146
Db	10681	aacttccaagatcttgaagaattcgcaataaagtcataaagcttgaaacaa	ca	10740
Qy	1147	aaaaaaatcacctttaccttctttaaagctcaagcatctgcagaaatcgag	atgcttgg	1206
Db	10741	aaaaaaatcacctttaccttctttaaagctcaagcatctgcagaaatcgag	atgcttgg	10794
Qy	1207	gccttctctctcctaactgtgttcttgagcaactacagctctccgacctc	aggagactct	1266
Db	10795	gccttctctctcctaactgtgttcttgagcaactacagctctccgacctc	aggagactct	10854
Qy	1267	agctctggctcagactcggataccttcgagaacagttccagtcacagctc	taagtcggg	1326
Db	10855	agctctggctcagactcggataccttcgagaacagttccagtcacagctc	taagtcggg	10914
Qy	1327	ctttatactgataagaactctttcgattactaacatcacagaaattatcga	aat	1386
Db	10915	ctttatactgataagaactctttcgattactaacatcacagaaattatcga	aat	10974
Qy	1387	aacaaagcgacagatgttgagggtgtgtcttacgtataaaggaaacctt	acttgtgaaaac	1446
Db	10975	aacaaagcgacagatgttgagggtgtgtcttacgtataaaggaaacctt	acttgtgaaaac	11034
Qy	1447	tctcacgcgtctacaatttttgaaaactcttcccgataaaacagtgaggga	aatctacgga	1506
Db	11035	tctcacgcgtctacaattttttgaaaactcttcccgataaaacagtgaggga	aatctacgga	11094
Qy	1507	gaacacaacatcacctctaatttgacaggggaagactctattccaagaga	aatctacgcc	1566
Db	11095	gaacacaacatcacctctaatttgacaggggaagactctattccaagaga	aatctacgcc	11154
Qy	1567	aaagaagaggcgggtgggactctctataaagggtacagataaagctcttc	acaatgacagga	1626
Db	11155	aaagaagaggcgggtgggactctctataaagggtacagataaagctcttc	acaatgacagga	11214
Qy	1627	ctggatagtcttctgtttaattaatacacatcagaaaacatcgtgggtgga	gccttttgtt	1686
Db	11215	ctggatagtcttctgtttaattaataatacacatcagaaaacatcgtgggtgga	gccttttgtt	11274
Qy	1687	accaaaagaatctctcagacitacacctctgatgtgaaaacatctccaggga	aatcacgcct	1746
Db	11275	accaaaagaatctctcagacitacacctctgatgtgaaaacatctccaggga	aatcacgcct	11334

Qy	1747	gtacaatggtgaacagcttactactggaataaattacagagagtaattggtggaggcgtg	1806
Db	11335	GTACATGCTGAAACAGTCACTACTGGCAATAAATCTACAGGAGGTAATGTTGGAGCGCTG	11394
Qy	1807	tgtcaaaacgtcttgcccttactctaaacctcaaaagcatttctatatccggaattctgca	1866
Db	11395	TGTACAAACGCTTGGCTTATCTAACCTTCAAAGCATTTCTATATCGGGAATTCGTGCA	11454
Qy	1867	gcagaaaatggtgtggagcccacacatgcccagatagcttcccacagcgcgatactgca	1926
Db	11455	GCTGAAATGTTGGTGAGCCACACATGCCAGATAGCTTCCCAACGGCGATCTGCA	11514
Qy	1927	gaacgcccgcagcagcttctgcgcgcagcttactctccaaatctgcccc---ggctcga	1983
Db	11515	GAACGCCGCGACAGCTTCTGCCGCGACGCTACTCCCGAGTCTGCCACAGTGGTCTCA	11574
Qy	1984	actgctctaagcacaccttcaattcttaccgtctcttcaattaaaccttactagagcctct	2043
Db	11575	ACTGCTTAAGCACACCTTCATCTTACCCTCTCTTCATTAACTTACTTACGAGCCCT	11634
Qy	2044	tcaacgctctcctgcgaaccttaataaggaacctcaagatccttaatgctgtgatacagac	2103
Db	11635	TCACAAGCCTCTCTCGCAACCTCTAATAAGGAACCTCAAGATCTTAATGCTGTATACAG	11694
Qy	2104	ttattgatcgattatgtagtgtatgatacgactatcagcaaaaacactgctaaagaggcgt	2163
Db	11695	TTATTGATCGATTATGTAGTTGATACGACTATCAGCAAAAACACTGCTTAAGAAAGCGGT	11754
Qy	2164	ggaatctatgtataaaagccaagatgtccgcgcatagaccaaactgaatatctctgagaac	2223
Db	11755	GGAATCTATGCTAAAAAGCCAAAGATGTCCGGCATAGACCAACTGAATATCTCTGAGAAC	11814
Qy	2224	tccgctacagatagtgagggtatctgctgtataaagaatcttagaactagatgctcta	2283
Db	11815	TCCGCTACAGAGATAGGTGGAGGTATCTGCTGTAAAGAATCTTTAGAACTAGATGCCCTA	11874
Qy	2284	gtctccttatctgtaacagagaaacctgtgtggaaagaaggtggaggcttacatgctaaa	2343
Db	11875	GTCCTCTTATCTGTAACAGAGAACCTTGTGGAAAGAAGGTGGAGGCTTACATGCTAAA	11934
Qy	2344	actgtaaatatattctaattctgaaatcaggcttctcttctcgacacaaagcaaaactcc	2403
Db	11935	ACTGTAAATATTTCTTAATCTGAATCAGGCTCTCTTTCTCGAACAAACAGCAAACTCC	11994
Qy	2404	tcattccagagctgcgaacaacagcttcagaccctctgcagctctgctgtcctccctacaa	2463
Db	11995	TCATCCACAGAGTGCACAACACAGCTTCAGACCTGCTGCAGCTGCTGCTCCCTACAAA	12054
Qy	2464	gcagccgcagcgcgcaccatcatctccagcaacacaaacttatcagggtgtagtagga	2523
Db	12055	GCAGCCGCAGACGGTACCATCATCTCCAGCAACACCACTTATTACGGTGTAGTAGGA	12114
Qy	2524	ggagctatctatggaaaaaggtttacattctctcaatgacgggaactgtcagttctct	2583
Db	12115	GGAGCTATCTATGGAAAAAGGTTACATTTCTCTCAATGTACGGGACNTGTCAGTCTCT	12174
Qy	2584	gggaaccaagctatcgataaacaatccctcccaatcatcgttgaacgtacagaggagacc	2643
Db	12175	GGGAACCAAGCTATCGATAACAANTCCCTCCCAATCATCTGTGAACGTACAAGGAGGCC	12234
Qy	2644	atctatgcgcaaaacctttgtctattggattcttcgatgctggaaacctctatatttc	2703
Db	12235	ATCTATGCAAAACCTCTTGTCTATTGGATCTTCCGATGTGGAACTCCTATATTTTC	12294
Qy	2704	tgggggaacagtgtctccactgggaaaacctcaaacacagggcaaatagcgggaggcgt	2763
Db	12295	TCGGGGAACAGTGTCTCCACTGGGAAATCTCAAAACAGGGGCAAAATAGCGGGAGGCG	12354
Qy	2764	atctactcccactgtttacattgaattgaattgctcgacattctctaaacaatacacacctct	2823
Db	12355	ATCTACTCCCCTACTGTGTACATTTGAATTTGCTCGGACATCTCTTAACAAATACAGCCCT	12414
Qy	2824	atagctacacgaagaacttcttgaagaatggatcctcaggaataattctatlaagataacc	2883





Qy	5044	tctatctacagaaataacccctgtctctgtataatcggtggtattgtcttcttcgaatgaagctggt	5103
Db	14635	TCTATCTACAGAAATAACCTGTCTGTAAATATCGGTATGTTCTTCCGAATCAAGCTGGT	14694
Qy	5104	caagtattctcggagtgccaaactagacacctctctagagcagatacagctactcaacta	5163
Db	14695	CAAGTTATCTCGGAGTGCCAACTAGAACCTCTCTGTAGAGCAGATACTACTCAACTA	14754
Qy	5164	tatcttggtccctctcgtgactctctacagaaactatactatcgtatgtaggcattatcag	5223
Db	14755	TATCTTGTCCTCTCTGACTCTCTACGGAACATATATCTATCGATGTAGGCATCTATACG	14814
Qy	5224	ctatcgcaaatgactagctcgggtgctgcagatctctctaa	5265
Db	14815	CTATCGCAAAATGACTAGCTCGGCTCGCATGATCTCTCAA	14856
RESULT	2		
LOCUS	AE002338	12838 bp	DNA BCT 26-MAY-2000
DEFINITION	Chlamydia muridarum, section 66 of 85 of the complete genome.		
ACCESSION	AE002338	AE002160	
VERSION	AE002338.2	GI:8163293	
KEYWORDS			
SOURCE	Chlamydia muridarum.		
ORGANISM	Chlamydia muridarum.		
REFERENCE	1 (bases 1 to 12838)		
AUTHORS	Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.		
TITLE	Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39		
JOURNAL	Nucleic Acids Res. 28 (6), 1397-1406 (2000)		
MEDLINE	20150255		
PUBMED	10684935		
REFERENCE	2 (bases 1 to 12838)		
AUTHORS	Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	On Jun 1, 2000 this sequence version replaced gi:7190724.		
FEATURES	Location/Qualifiers		
source	1. 12838		
	/organism="Chlamydia muridarum"		
	/db_xref="taxon:83560"		
	/note="synonym: Chlamydia trachomatis MoPn"		
gene	69. .3029		
CDS	/gene="TC0693"		
	/note="similar to GB:AE001363; identified by sequence similarity; putative"		
	/codon_start=1		
	/transl_table=11		
	/product="polymorphic membrane protein A family"		
	/protein_id="AAF39509.1"		
	/db_xref="GI:7190725"		
	/translation="MPQALLIKELMQVITKIALCYQKYISRASNKTFTSHNTLSLSL LKCLGLLIYTSFAPGEMELAIISGHYKDRDAFTMISSCPEGNTYMINRKLIISE FSLDKFSSGAGFNLAGKIAFLGKSSSIIFKHLNMGFGSGIFSESIAIEFSDLRK LVAFSSGTGIFTARDIDISKNNHYIAFRNNIAKNGVILLQGDGRGVSTDDQG LIAFNNQALVSPSIKHSRGGAISGDFAGSRIFLNNQOITFEENSVAHGALYAKN GVVEFLNGGTLSPKENSTRANGAIYTGFRKANQQTAPIIFSONNANQKGALYAOY VNI EQNDAIRFENNSAKGGGAISSQCAITAHNSITFFSNFAGDIGGGAILLGGQ PSLSLVAHNGNIAPIGNTMLPATKASLPNNNSILVKESPYKTOFAANKNRSIIFDP		

gene  
CDSgene  
CDS

VIALSPSTSPVEINSPHEPFPFKGTIVFSGANLIDDAKEDIANRTSIFNOPVLHL  
NGTSLBEGAHVVOQPKOTGGRISSLPSSGLALYTTNTLFGHVSNDPDLLEINGLSL  
GVDTSPNLYSEIRAGSAPIKLSQSDHPERLFYENRDSAAAPYOMEILLSDKII  
DVSNFTIDAPVKNKEAGFOGSHFSPQNPVNTVHKVLRASWPTGEYILEPSRVGN  
AIPNSUMSTELLQITASHNLGDHLCNNDSPTYSYLLGLLGGIAEMRTYSKESFI  
SRGTGTGTTIIRLTPTTLTSGGATHMFGDSFVTKLPEFIASEGVMQVNGITQILGPT  
VKSTLCAALDHAMIRLSAQNHRTKWDFTGIRIGLGSYSLIDYENMVRIFTFANV  
EATNVLOKSTETGYNPRSFARLTNIAVPVGIYGEFCLSNHSFALLGKHGIGYSRD  
IKRNPVTFQAOLAMNFSWTANGCQVPSAHTANQLILLRYKACSLYVNAVATKLEST  
YLSSSUSCGGYVGF"

3138. .8176  
/gene="TC0694"  
3158. .8176  
/gene="TC0694"  
/note="similar to GP:4376830; identified by sequence similarity; putative"

/codon\_start=1  
/transl\_table=11  
/product="polymorphic membrane protein B/C family"  
/protein\_id="AAF39510.1"  
/db\_xref="GI:7190726"

/translation="MSSMKWLSATAFAAVLPVSVSGRCFPESKELNFSRTGTSSTTF  
TETIENGTEYIVSGSSFTNIPVKKPTDDSTSTPTTSVADPTBKIVRASS  
SSPNSGDSATPOPKGGGAFYNEHSGILSFMAKSGVEGSLTLSNKKMGCGAIYSQG  
ELLFTDLTGLTIQNLSQLSGGIFGGSTISFSGINQATFSNTAEVVPBETTPNPNP  
GTOTTTSPSPSTKVSQSLFTYSSSTOANGADSOPTSHKPGGGAIYATGDLTISDS  
OEIVFSYNKASKDGGAIYFAEKVYSENIITTLKVQNGAEKGGIYASGDLSTQSSQK  
SLFNSNISKOGGALYIEGNVDFKDLSEIRIKYNKSGTFEKKVTLSEPAQNTKSSV  
TAASQSGPNTPTPTPVTAKGGLYTEKNLSISNITGIIETNNKATDVGGAIVKVG  
DKTLTKDSHRLQFKNSKKGGLYEDTITLSNLTGTLTQENTKEEGGLYIQDG  
KATGTGDSFCLIDNTSATHGGAIVTKEISQYTSVDEVEFPGIPVHGETIISGN  
KATGSGGGVCTKHLVLSNLTISSENFASENGGACTCPDNPAPAPTPTSTNOTA  
APDKDKFLDYVVSVTIDNKATKACGAYAKKLSRIDELNISNAQVGGGFC  
CTSELDITIASLVTKNLAGKGGGHLAKTLNISNLSGLSFNSNNTASSTGCVATT  
CYSQSPSTVSSFLPRATAGSPAPQPTIYAVVGGAIYGETVFSKCSGLCQATTEN  
SAIDNTPSPNLQVGGAIYSLTSEADPSTSYVSKNSVSTGRTAQKTGTAGGA  
IYSPSVTLNCTQTVFSGNSAMATNPPGTSKPDITGGAIGTAGTISLSTKSHFSENTA  
DLGAAIGTLSGGSSNLTERITLSNGSFTFEKNKANKRGVIYAPSVSIKGNITFNON  
TSTHDGSAIYETKDATIESGLSVLFTGNVTAQASSTAGTGGOTTNTNYCAAIEGDP  
GTOTDTTLALIASGNITFSNNSQNTATNNPATKFCISGYKLTILQAAGKTIISFF  
DSIRTSKTKGQAQNSYETIDINKTNSYTAGTVLFSSELHELKSVKSPQNVVJHNTG  
LVLKNAELHVSEFEQKESGLIMEPAGVLSNQINANGALAINGLTIDLSLGAPQTC  
EVSPPELRIVATTSNGGGVGYVTASKNLSAASPVTAAATNPTMADKNKVTGTAL  
TLDPDGNFYQNPILGDLTDVPLIKLPTTANQVDSNLTLSGLDSPKPKVGTWTINL  
PDQGTQKVVANWKFDMYRWEYIPRDNHEVANSILGSONMIVVKOGLINNLHNAHF  
DDAAYNNFVSYGYCTELSQOGLTSEEFYSYSGRTSVAIDAKPRDPIILGAESKMWG  
RTAKKKNVNSHKGSEYTSQASVYGKFLIFLLNKHQWALPFLQGVVSYGHKHD  
TTLYPSIHEKNKGDWEDLGLVLRVSMQVKEPSKRKRKVALYGELEYSYTRQKSF  
TEIDYDPRFDDCAIRNLSPMGCYFEGAIMSYDILMYNKLISLAYMPSIYRNPPVKY  
WLSNSETSKVWCVPTRTSARAEBYSTQLYLGPEWTLYGNTYIDVGYTLAQMTSCGA  
RMIF"

8355. .12737  
/gene="TC0695"  
8355. .12737  
/gene="TC0695"  
/note="similar to GP:4376830; identified by sequence similarity; putative"

/codon\_start=1  
/transl\_table=11  
/product="polymorphic membrane protein B/C family"  
/protein\_id="AAF39511.1"  
/db\_xref="GI:7190727"

/translation="MKFLSATAVFAAALPSITSASSVESOIETKDLNLSRTGSSSSOS  
FTEILPENGAEYRVSGDYSDFSNIPEEAEETLAIHKEQPNNEVLSSENHQASFOQ  
SAQNTENASEGNSPNSNENQNSSTETESITTDQVQNDNESAASTVPTIVETATAMR  
LPYHQTESLVEGATEEDQNQNSQTSSSGAFYNSQDPSLFINDPDKDSTLTLIS  
TVRPPAAKPELPNDPNSGNSDGSNSGNTDNEPNNSASCAESTISISSYDILTSFN  
KIVYIGSGAIYKGPLSTGLKALKENLSOKAGGAIKAEISTISISSYDILTSFN  
TPSAOLPNPATPFLSYSTNSQPTDTEPENAWHAESGGGAIYKSKGLSTASKEVVF  
DHNSATKNGAIGEEBIALEKLTASKEFNTGEGKGAIAHAKTYLSDASIKNTLIFYN  
NTAKTPENSLSQLANQNPSEEHQDTESEGESQLETSPTINQDSASVHAIRFS  
IAASSGQSTENPNADGTSAGDSSQSTPPGSDSIHNVKLSLAYMPSIYRNPPVKY  
NLSSGCTSNNAVDHQISGSTSDVLGGAIYAKTSLTIDSGNSSGTTTFSNTTSSKS

TTTCQVAGGAIFSPSVITTTTPVTPSKNSAINATTSKKDTPFGGALGALSIVTSLSKGARF  
SENIAQSGAIGLVPITQADATVQLTGTSGYIFPKNKALKRATVAPISVTSKAHTATFD  
QNISABEGSAIVTEKATIESLGSYDITGLNLTPIQISTVLTSGNTSKYAAALFGGIA  
NAGSDOTRLPTKLGLASAGYISPRNNEYRNDATGTQSCFSTAGDIKLTMOAAEGKV  
ISFFDAITSTKTKTGTSLANDYDITLKNKSDGSIINSAFTGTIMFSESSELHENSXIYIQ  
NVYLSGSLLIKANTELHVLSFDQKEGSSLLIMEPCSVLSNQDITADGSLVYNSUTIDLS  
SVGRNSAGDNIIMPFLERIVDTSTNSGSSSTPPSSNTTPPSTPTAQAPISKNFAAT  
TTPVNTGTNGIVNGLIKIDTPNGTFTQNPALGSDQKISLLVLPDQTKLQAQKV  
DITGDSPPKRGYTGTLTDLPQQLQVYIQALMTFKSYROWAYIPRONHFVYISGLAQ  
MSMATVKQGLINDKLNAREDEYAVNLMTSGLGTLMSORGORSREMTYVYRSGA  
LDAKPTQDILLIGAFAKSMITGRSKSLERNYTHKSGSEYVQYASVYGGSPFYLTINKEA  
GRSLPULLQGSYGIYKHTDVTHTPIRELNKGWEDLGMVLTALRKSLLTKTPKQGD  
SKRITVYGEVYESSIKQKFTETEDYDPFVSNCTYRNLYAPALLEGEGKNGDILMY  
NRESVAYMPISYIRNSPVCKYVLLSGEGGIEVCGVPTNRSSRAEYSTQLYVLGLPWLTLY  
GSYTIPADAHTIANMTNCGARMT<sup>24</sup>



Db 6887 G---TTGCGAGCTACAAATCCAACTGCGAGACAAAGAGTCTTCTTAACCGCGCTCTG 6943  
QY actttaataagatcctaataagaaacttttaccaaaacccctatgttaggaagagcgtacta--- 4092  
Db 6944 ACTTTGATCGATCCAGATGGAACCTTTTACCAGAAATCCTATCTTAGGGACAGATTTAACC 7003  
QY gatgtaccactaataaagttcccgactaacacaaagtgacgtccaaagtcctatgatttaact 4152  
Db 7004 GATGTACCGCTTAATCAAACTCCCTACACAGCAAAATCAAGTTGATGTTTCTAACTTAACA 7063  
QY ttatctgggagatcttttccctcagaagaaggtacatggaacctggaacattagattctaact 4212  
Db 7064 CTATCCGGAGACCTTTCTCTAGAAGGGGTATACAGGAACCTTGGACGCTAAATCTCTGAT 7123  
QY ccacaaacaggggaaacttcaagccagatggacattcgatcactctatcgctcggtggtatc 4272  
Db 7124 CCACAAACAGGAAAGTTGTCGCAACTGGAAATTCGACATGTACCGCGTTGGGAATAC 7183  
QY atacctaggagataatacttttatcgaaactctatcttaggtctcccaaaactcaatgatt 4332  
Db 7184 ATACCAAGAGATAACCATTTCTATGCAATTCGATCTGGGCTCCCAAACTCCCATGATT 7243  
QY gttgtgaagcaagggcttatcaacaacatgttgaataaagcccgcttgcgatgatcgcgt 4392  
Db 7244 GTTGTGAACAAGGGCTTATTAATAACATGTTACATAATGCCCGTTTCGATGACGCTGCT 7303  
QY tacaataactctgggttcagagtaggaactttcttagctcaacaagaaactcctct 4452  
Db 7304 TACACAACTCTGGGCTCAGAGTGGAACTTCCATTCGCAACAAAGTACTCCTCTTT 7363  
QY tccgaagaattcagttactacagcgcggaacttcagttggccatcgatgccaaacctaga 4512  
Db 7364 TCTGAAGAGCTTAGTTATTATAGCCGAGGAACCTCTGTAGCTATCGATGCAAAACCCACGA 7423  
QY caagattttatcctaggagactgatttagtaagatagtagtggggaagaaacccatcaaaa 4572  
Db 7424 CCAGATTTTATCCTGGAGCTGCTTTAGTAAGATGGTAGGAAGAACTAAAGCGATTAAG 7483  
QY aaaaatgataaataacttccatcgaaggtctgagtagctcttaccagaactctctgctatgga 4632  
Db 7484 AAGGTACATAAATTATCCCATAAGGATCCGAATACTCATACCAACGCTCTGTTTACGGA 7543  
QY gbtaaattcctgtattcttctgtcctaataagaacacatggttgggcaacttcttctcctaata 4692  
Db 7544 GGTAAAGTCTCTGACTCTTACTTAATAAACAACATGCGTGGGCACACTACCTCTCTTATTA 7603  
QY caagagatcgctcctatggacatatataaacaatgatacaacaacactttacccttctatc 4752  
Db 7604 CAAGGGGTGCTATCCPATGGGCATATAAGCATGATACAAACACCTCTATCCTTCTATC 7663  
QY catgaagaataaagagagattgggaagatttaggatgggttagcggtatcttgcgtatctct 4812  
Db 7664 CATGAAAAAATAAAGGACACTGGGAAGATCTGGTGTGGTTAGTAGATCTACCGTCTCT 7723  
QY atgattcttaagaacacttctaagaattcttctaacacgatacactgtctatggggaactc 4872  
Db 7724 ATGGATGTTAAAGAACCTTCCAAGAGATCTTCTAAACCGGTTGCTTTACGGAGAAT 7783  
QY gagtattccagacttgcgcagaacacagttcacagaataatcgattacagatccaaagacattc 4932  
Db 7784 GAGTACTCAGCATTCGGTCAGAAATCATTCACGGAAATTTGACTACGATCCTCGACGCTTT 7843  
QY gatgattgcttaccagaatactctgcgtctcctgtgggagtcgctgcgtcgaagagactc 4992  
Db 7844 GATGATTGCTTACAGAAACCTATCCATTCCTATGGCTGCTATTTTGAAGGGGTATC 7903  
QY atgaactgtaattatttattgtataataagcttggaattagcctgacatgccttctatctac 5052  
Db 7904 ATGAGTTATGATATCTCATGTATATAACAAGCTTTCCTTAGCCTATATGCGCATCTATTAC 7963  
QY agaaataactcctgtctgtataatcggggtattgtcttcgaatgaagctgggtcaagtattc 5112

Db 7964 AGAAATAACCCCTGCTCTGTAAATATTGGGTCTCTATCTCGAATGAACACTAGCAAAAGTTGTC 8023  
QY tgcggagtgccaaactagaacctctgtagagcaagaatacagtagtactcaactatcttgggt 5172  
Db 8024 TGTGGAGTCCCCACAAACATCCGCTCGAGCAGAGATATAGTACACAACCTTTATCTAGGG 8083  
QY ccccttggaactctctacgaaatactactatcgatgtaggcattgtagcgtatcgatcgcaa 5232  
Db 8084 CCTTCTGGACCTTATATGGGAATTACCATAGACGTGGTATGTACACATTGGCACAA 8143  
QY atgactagctcggtgctcgcatgactctctaa 5265  
Db 8144 ATGACGAGTTGTCGCCCTATGATTTTCTTAA 8176

RESULT 3  
AE001315  
LOCUS AE001315 10199 bp DNA BCT 30-OCT-2000  
DEFINITION Chlamydia trachomatis section 42 of 87 of the complete genome.  
ACCESSION AE001315 AE001273  
VERSION AE001315.1 GI:3328842  
KEYWORDS  
SOURCE Chlamydia trachomatis.  
ORGANISM Chlamydia trachomatis.  
REFERENCE 1 (bases 1 to 10199)  
AUTHORS Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R., Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V. and Davis,R.W.  
TITLE Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
JOURNAL Science 282 (5389), 754-759 (1998)  
MEDLINE 9900809  
PUBMED 9784136  
REFERENCE 2 (bases 1 to 10199)  
AUTHORS Kalman,S., Mitchell,W., Marathe,R., Lammel,C.J., Fan,J., Hyman,R.W., Olinger,L., Grimwood,J., Davis,R.W. and Stephens,R.S.  
TITLE Comparative genomes of Chlamydia pneumoniae and C. trachomatis  
JOURNAL Nat. Genet. 21 (4), 385-389 (1999)  
MEDLINE 9920606  
PUBMED 10192388  
REFERENCE 3 (bases 1 to 10199)  
AUTHORS Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R., Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V. and Davis,R.W.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAY-1998) Program in Infectious Diseases, University of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA  
FEATURES  
source  
location/Qualifiers  
1..10199  
/organism="Chlamydia trachomatis"  
/strain="D/UW-3/Cx"  
/db\_xref="taxon:813"  
120..5432  
/gene="pmpc"  
/note="CT414"  
120..5432  
/gene="pmpc"  
/codon\_start=1  
/transl\_table=11  
/product="Putative outer membrane protein C"  
/protein\_id="AAC68011.1"  
/db\_xref="GI:3328843"  
/translation="MKFSATAVFAAALSSVTEASSIODIKNTDCNVSKLGVSTSOA  
FTDMLADNTEYRAADSVFYDFSTSSRLPRKHLSSSEASPTTGGVSSSSSGEYDEK  
TYGEKEVFENIKSLLEVNIATVEKGVSYAKERVSLNENVTAEATSSNGEGGGGIY  
SQDMLISDCNNVHFQNAAGATAVKQCLDEEMIVLLAEVDLSLSDTLDSTPEEQY  
ESNGODGSSEDTQVSESPSTPSDDVLKGKGGIVTEKSLTITGTTIDFVSN  
ATDSGAGVFTKENLSCTNTNSLOFLKNSAGOHGGAYVOTMTSVNTTSSTSTPLI  
GEVISENTAKHGCGGICTNKLSLNLTVTTLTKSKESGGAIFTDLASITPTPE  
STPSSSPASTPEVVASAKINRFFASTAKPAAPSLTEAESDQTDQTDFTSDNSDIDV  
STENILINAINQNTSAKKGATYGRKAKLSRLNNLELSGNSSQDVGGLCLTSEVFEF

AIGSLSHYNSAAKEGAIHSKVTLSNLKSTFTFADNTVKAIVESTPEAPEEIPVPE
GEESTATDPSNTEGSSANTNLGSGQDADTGTGDVNNESODTSDTGNAESEQLQ
DSTOSNEENTLPNSIDQSNENTDESSDSTEETITDESVSSESSTPDQGGAASS
GAPGSDQSIANAACLAKSYAASDSDSPVSNSSGSEEPVTSSSDSVDYASDNPDSSSS
SSAGDSEEPPEAGSTETLTLLIGGAIYGETVKIENFSGGIFSGNKAIDNTTGG
SSKSDVLGAAYAKTFLNLDSSSRRTVTFSGNTVSSQDTGQVAGAIYSPTVTIA
TPVFSKNSTNANANTTDQRKDTFGGAIGATSAVLSGGAHLENVADLIGSAIGLV
PGTONTETVLESQSYFPEKNALKRATIVAPVVISIKAYTATFNQNSLEEGSAIYFT
KEASIESGVLFTGNLVTLTSTTTEGTPATSGDVTKYCAALFGQIASSNGSQTDN
LPLKLASGNCIFERNNEYRPTSDTGTSTFCSTAGDVKLTMAAKAKTISFFDAIT
STKTKTQATAYDPLDINKSEDSVTSNAFTGTLFSESELHENKSIIPQNVLHSGSL
VLKNTPELHWISYFEKQSGSLMTPGSLVSNQTVADGALVINNMKTIDLSNVEKNGIAE
GNIPFTPELRITDITTCGSGGTPTSDSESNQSDTEEQNNNDASNGESSSPA
VAAHTSRTNRNFAAATATPTTPTATPTTTTNOVILGGEIKLIDPNCGTFFQNPAIRD
QOISLLVPLDSSKMAQKIVLJGDIAPOKGYTCTLDPDQLONGTISVLWKPDSYR
QWAVYPRDNHFIANSILGSOMLVTVYKQGLLNDKMNLARFEVSYNNLISGLGTMLS
QVGTPTSEEFYTSRGAVALDAKPAHDVIVGAASFVKMIGKTKSLKRENNYTHKGSY
SYQASVYGGKPFHEVINRKTEKSLPLLOGVISYGIKHDTVTHYPTIRERNKEWED
LGMWLTALRVSVLTPAQGDTKRITVYGELEYSSIRQKQETETEDPRYDNCITYRL
AIPMLAFEGELSGNDILMYNRFSVAYMLSIYRNSPTCKYQVLSGEGGEIICGVPTR
NSARGEYSTOLYLGPLATLYGXYTIEADAHTLAHMMNCGARMTF"
complement(5593..5676)
/gene="tRNAser\_3"
/product="tRNA-Ser"
complement(5593..5676)
5985..6815
/gene="yebl"
/note="CT415"
5985..6815
/gene="yebl"
/codon\_start=1
/transl\_table=11
/product="Solute-binding protein"
/protein\_id="AAC68012.1"
/db\_xref="GI:3328844"
/translacion="MRLFLILLFSLGITCSYGDVSTRKQILVSIVPKFLVEQIAGD
TCQVTSVMONDPHDSPHTELPKYIEKIQVELWFKIGEGETKCTERIIISCKVDLAAN
IDKFTSAGACORFLSDFTHTWLSPKNLKIQIOALTEALVETAPHEHTYRKNCSLOS
OLDLLDKQISIVSSTSORWLVTHGAFAYFCRDYGFIOHTIERANSELSPKDVYRV
ERTIRDNHLSVILLKHAGKSSAALVRKFNMTPIILDPAEDVFNLLALATAFANL
"
6812..7522
/gene="CT416"
6812..7522
/gene="CT416"
/codon\_start=1
/transl\_table=11
/product="ABC Transporter ATPase"
/protein\_id="AAC68013.1"
/db\_xref="GI:3328845"
/translacion="MTKQMLLENVSFRYKGTGLIVDHVSCEVYSGDFIGIIPNGGG
KTTTLQMLGLQPCVSGISYTSVDHRPLUSIGWVPOHFSYDAAPFITVKTVELUSGR
ATLWYGRYTODEHEAAEALLVDLIDYKDSFSLHSGGQIQRIORVILARALAPPEL
LLEPTANIDPVNQKTLQILSALNKHCTILMITHDLHHTAGCFNRFVFNKKTLLTILA
DTTTISRFCNTEGRCp"
7513..8394
/gene="CT417"
7513..8394
/gene="CT417"
/codon\_start=1
/transl\_table=11
/product="(Metal Transport Protein)"
/protein\_id="AAC68014.1"
/db\_xref="GI:3328846"
/translacion="MPMISILCSLFPPLPFLSLAAGFASIAAGIIGSIYVVKRIVSI
SGWSTHGLTGQIALWQFNLPISPLHGAIASAFVAICIGNVILKHYERDSDSI
SMWSIGMAIGICIGIKLPSFNSLSDFGNILMWTPODLYFLGILDFTVATVSI
HTRFLALCFDEKYNALHYSIKTWYLLLLLTALTTVVLMYVMGVILMLMLPVS
ACRSYKWSHIIYIASILNIVCSFLGIMLAYLLDLPVGTAILMGGAYSLSLLKRS
YNASTPSPVSPESKINS"
complement(8310..9317)
/gene="yhb2"

/note="CT418"
complement(8310..9317)
/gene="yhb2"
/codon\_start=1
/transl\_table=11
/product="GTP Binding Protein"
/protein\_id="AAC68015.1"
/db\_xref="GI:3328847"
/translacion="MFYDQITLLELRAGKNGVVAWRKEKYLPKGPGYGGNGGSGSI
LIEATNMYSEYRNRLFKADQAGASNNRTGRNGKDLVLKVPSTGILLRDNATGE
LIHDTFKDGERIVVCGRGKGNVFKFTSTNRAPTKATPKGPEGILVELEKLAD
IGLFTFNAGCKSTLFTLARTEVKVGAYPTTTLHPSGLVHQEGMLQKPMWADIFG
IEGASNRGIGLDLRIERTRLLEFVIDISGIERHSPQDLKILMAGELLAYKEELK
DKDMVIALNKIDQLLPDERERVALLKQFPDQFEILLSLGTGSGVDALYDLPKSKLS
E"
complement(9407..9658)
/gene="rl27"
/note="CT419"
complement(9407..9658)
/gene="rl27"
/codon\_start=1
/transl\_table=11
/product="L27 Ribosomal Protein"
/protein\_id="AAC68016.1"
/db\_xref="GI:3328848"
/translacion="MAHKKGQASRGRDSESKRLGLKVGAGQVRVSTGSLVRQGRTK
WHPAVNVGRGKDDTLFALVDGIVVMKKTDRITYYSVIPQA"
complement(9689..10012)
/gene="rl21"
/note="CT420"
complement(9689..10012)
/gene="rl21"
/codon\_start=1
/transl\_table=11
/product="L21 Ribosomal Protein"
/protein\_id="AAC68017.1"
/db\_xref="GI:3328849"
/translacion="MEPYAVIOTGNKOYQVRKGDVIDVELLDGISEENKEVLFQDVLV
TFQGEKASVGAPTVGNVVKGLSVFYVRGEKVWAYKYKKRKNYHKIGHRONLYLRVKI
SDLVW"
BASE COUNT 3079 a 2313 c 1934 g 2873 t
ORIGIN
Query Match 15.7%; Score 824.2; DB 1; Length 10199;
Best Local Similarity 59.0%; Pred. No. 2.8e-166;
Matches 1765; Conservative 0; Mismatches 1018; Indels 210; Gaps 12;
Qy 2426 cagcttcagcacctgtcgcagctgctctccctacaagcagcgcgcagccat 2485
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2497 CAGACTCTTCCTCATCTGGAGATCGCTGGAGACTCTGAAGAACCGACTGAGCCAGAAG 2556
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2486 catctccagcaacacacattattcaggtgtagtaggaggagctatctatgagagaaagg 2545
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2557 CTGGTTCTACAAACAGAACTCTTACTTTAATAGGAGGAGGTCTATCTATGAGAAACTG 2616
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2546 ttacattctctaatgtagcgggactgttcagttctctcgggaacccaagctatcgataaca 2605
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2617 TTAGATTGAGAACTCTCTGCGCAAGGAATATTTTGGAAACAAAGCTATCGATAACA 2676
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2606 -----atccctcccaatcatcgttgaaacgtacaaggaggagccatctatgcacaaacct 2659
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2677 CCACAGAGAGGCTCCTCTTCCAAATCTGACGCTCTCGAGGTGCGGTCTATGCTATAACAT 2736
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2660 ctttgtctattggatctctccagatcgtcggaacctctctattttctcggggaacagtgct 2719
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2737 TGTTTAATCTCATAGCGGGAGCTCTAGACGAACCTGTACCTTCTCCGGAATACTGCT 2796
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2720 ccaactgggaaatctcaacaacagggcaaatagcgggagcgagctactcctactg 2779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2797 CTTC-----CAATCTACACAGAGCTCAGGTTCGAGGAGCTATCTACTCTCTACTG 2850
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2780 ttacattgaattgtcctcgcgacattctctaaacaatacacagcctctatagctacaccgaaga 2839
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||



Db	2851	TAACCAATTGCTACTCTCTGTAGTATTTCTTAAAAACTCTGCAACAAACAATGCTAATAACA	2910
Qy	2840	cttctctgaagatggatcctcaggaaattctctataaagataccattcggagagcgaattg	2899
Db	2911	CT-----ACAGATATCAGAGAAAGACACCTTTGGAGGAGCTATCG	2952
Qy	2900	caggga---cagccattaccctatctggagtcctcgtatttccagggaatacagcgtgatt	2956
Db	2953	GAGCTACTTCTGCTGTTCTCTATCAGGAGGGCTCATTTCTTAGAAAACGTTGCTGACC	3012
Qy	2957	taggagctgcaataggaaactctagctaatgcaaatataccagtcgaactagcggatctc	3016
Db	3013	TCGGATCTGCTATTGGGTGGTGCCAGGCACACA-----	3046
Qy	3017	aaaatagcattacagaaaaaattactttagaaaaacggtcttttatttttgaagaataacc	3076
Db	3047	-----AAATACAGAAACAGTGAATTAGAGTCTGGCTCTACTACTTCTGAAAAAATA	3099
Qy	3077	aagctaataacgtggagcgattactctctagcgtttccattaaagggaataataatta	3136
Db	3100	AAGCTTTAAAACGAGCTACTATTATCGCACCTGTCTGTTTCCATTAAAGCCTATATCTCGCA	3159
Qy	3137	cttccaatacaaatacatccactcctatgtaggaagcgctatctacttttcaaaaaagtgcta	3196
Db	3160	CATTTAACAAAAACAGATCTTAGAAGAAGGAGCGGATTTACTTTACAAAAGAACGAT	3219
Qy	3197	cgattgagcttttaggatctgtctttttcacaggaaataacggttacagctacacaagcta	3256
Db	3220	CTATTGAGCTTTTAGGCTCTGTCTCTTCACAGGAAACTTAGTAAACCTTACGGCTAAGCA	3279
Qy	3257	gtctgcacaatctggacaaaatacaaa-----atactgccaactatggggcagcca	3307
Db	3280	CAACTACAGAAGGCACACCAGCCACAACTCAGGAGATGTAACAAAATATGGTCTCTCTA	3339
Qy	3308	tctttggagatccaggaaacactcctaactcgtctcaaaacagatgcattt---taaccccttc	3364
Db	3340	TCTTTGGACAAATAGCAAGCTCAAAAGGATCTCAGACGGATAACTTCCCTCGAAATCA	3399
Qy	3365	tgccttctctggaacattacttttagcaacaacagtttacagaataaacaagggtgata	3424
Db	3400	TTGGTTCAGGAGGAATATTTGTTCCGAAACAATGAATACCGTCTACTTCTTCTGATA	3459
Qy	3425	ctccgcctagcaagttttgtagtattgcaggatacgtcaaaactctctcaagcgccta	3484
Db	3460	CCGGAACCTCTACTTCTGTAGTATTGGGGGAGATGTTAAATTAAACCATCAAGCTGCAA	3519
Qy	3485	aaggggaagactatagcttttctgatgtgtgtcacacctctaccaaaaaacagggttcaa	3544
Db	3520	AAGGGAAAACGATCAGTCTTCTTTGTATGCAATCCGGACCTCTACTAAGAAAACAGGTACAC	3579
Qy	3545	cacaaaacgtttatgaaacttttagatattataataagaagaaa-----cagta	3592
Db	3580	AGGCAACTGCCTACGATCTCGATATTAAATAATCGAGGATTCAGAACTGTAAACT	3639
Qy	3593	atccatatacaggaactattgtgtctctctctgaattacatgaatacaaatcttacaatcc	3652
Db	3640	CTGGGTTTACAGGAACGATCTGTTCTCTCTCGAATTACATGAATAAATCCTATATTC	3699
Qy	3653	cacagaatgcaatccttccaacacggaactttagttctttaaagagaacagaactccaag	3712
Db	3700	CACAAAACGTAGTTCTACACAGTGGATCTCTTGTATTGAAGCCAAAATACCGAGCTTCATG	3759
Qy	3713	tagtctcttttgacgaagaagggtctaaatattatggaaccggagcgtgtgttat	3772
Db	3760	TTATTTCTTTTGGACAGAAGAAGGCTCTTCCTCGTTATGACACCTGGGATCTGCTCTT	3819
Qy	3773	ctaaccaaaacatagctaaacgagctctagctatcaatgggttaacgatgatctttcca	3832
Db	3820	CGAACAGACTGTGTGCTGATGGAGCTTTGGTCAATAAATAACATGACCATGATTTATCCA	3879
Qy	3833	gtatgggga-----ctctccaagcgggggaattctctctccccaagattacgtatcg	3886





QY	4360	atgttgtaataatgcccccttcgcatgatgatgatactgcttacaataaactcttgggtttccaggagta	4419
Db	4279	ATGTTGAACAATCAAGAGTTTGAAGATCTCGCTTTTCAACAACCTCTGGGCTTCGGCTATA	4338
Qy	4420	ggaactcttagctcaacaagaactcctcttcctcgaagaactcagtttaactacacgcgc	4479
Db	4339	GGATCTTCTCTTAGGAAGAAGATATCTCGAANAATCTGATCAATACCTATCATGGCGAGA	4398
Qy	4480	ggaacttcagttgcatcgatgcgaacactagacaagattttatctcctagagactgcattt	4539
Db	4399	GGCTATACCGCTGCTGTGGATGCCAAACCTCGCGCAAGAATTTATTTAGGACCTGCCCTC	4458
Qy	4540	agtaagatagtggggaaacccaaagccatacaaaaaaatgcataaataacttccataagggc	4599
Db	4459	AGTCAGGCTTTTGGTCACGCCGAGCTCTGAATATACCTTTGACAACCTATTAAGCATAAAGGC	4518
Qy	4600	tctgagtactcttaccaaagcttctctctctgaggttaacttctctgatttcttcttctcaat	4659
Db	4519	TCAGGTCACTCTACACAAAGCATCTCTTTATGTGGCAATATCTCTATTTTCTCTCGCGATA	4578
Qy	4660	aagcaaatggttgggcaacttcttccctaatacaagaagtgctgctctatggacattt	4719
Db	4579	C-----GGTCTCGGCTATTCTATTCACAGGTTGCGGACCTATGTTATATG	4626
Qy	4720	aaacatgatacaacaacttaccctctctatccatgaagaataaagagagatgggaa	4779
Db	4627	CAACATGACACCAACCTACTACTCTCTTATTTGAAGAAAAAATATGCAAACTGGGAT	4686
Qy	4780	gatttagatggttagcggatctcgtatctctatgatcttaagaactcttcaaat	4839
Db	4687	AGCAATCTCTGGTATTTGATCTGCGTTTCAGTGTGGATCTTAAGAAACCTCAACCTCAC	4746
Qy	4840	tcttcaaacggatcaactgtctatggggaactcgatattccagcattccgcagaaaacag	4899
Db	4747	TCTACAGCAAGCTTACCTTCTATACAGAAGCTGAGTATACCAGAATTCGCCAGGAGAA	4806
Qy	4900	tccagaaatcgattacgacatcccaagacacttcgatgttgcttaacagaatactgtcg	4959
Db	4807	TTCACAGAGCTAGACTATGATCTCTGATCTTCTCTGCTGCTCTTATGGAACCTTAGCA	4866
Qy	4960	cttccgtggaatgcgctcgaagagatcatatgaactgaattattcttctgtataat	5019
Db	4867	ATTCTACTGGATCTCTGAGACGGAGCATTAGCTGGCGTGAGATATTCTATATAT	4926
Qy	5020	aagcttgcattagcctacatgccttctatctacagaaataatcctgtctgtaaatatcgg	5079
Db	4927	AAAGTATACGCTGGTACCTCCCTGTGATCTCAGGAATAATCCAAACGACCTATGAA	4986
Qy	5080	gtattgtcttcgaatgaagctggctgaagtattatcgcggagtgcccaactagaacctctgct	5139
Db	4987	GTTCTCTCTACAAAGAAAGGCAACGCTAGTCAACGTTCTCCCTACAAAGAAACGAGCT	5046
Qy	5140	agacagaaatacagtagtactcaactatacttctgctcctcttgactcttccaggaactat	5199
Db	5047	CGTCAGAGGTGAGCTCTCAAAATTTATCTTGTGAAGTATCTGGACACTCTACCGCACGAT	5106
Qy	5200	actatcgatgtaggcattgatacgcctatcgcaaatgaactagctgcggtgctgcgatgac	5259
Db	5107	ACTATTGATCTTCATGATATCTTTAGTGCATAATGCCCAACGGAGGATCCGGTTTGTA	5166
Qy	5260	ttcta 5264	
Db	5167	TTCTA 5171	
RESULT	5		
LOCUS	AE001638		
DEFINITION	Chlamydia pneumoniae section 54 of 103 of the complete genome.		
ACCESSION	AE001638		
VERSION	AE001638.1		
KEYWORDS	14781 bp DNA BCT 01-DEC-2000		











```
QY 4180 gggatcatgggaacacctggacattagattctaatccacaacaggggaacacttcaagccaga 4239
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24912 GGATATTTAGGAACCTGGAAATTTGGATCCAAATTCCTCGGGTCAAAATTTATCTTAAAA 24971

QY 4240 tggacattcgataccatctatcgctgggtatcacatcaccttaggataataatcttttatgcg 4299
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24972 TGGACCTTTGACAAATACCTGCGCTGAGCCTACATCCCTAGAGACAACACATCTTACATC 25031

QY 4300 aactctatttaggctcccaaaactcaatgattgttggaagcaaggccttatcaacaac 4359
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25032 AACTCTATTGGGGAGACACAAACCTTTTAGTCACTGTGAACACAGGGATCTTAGGGAAC 25091

QY 4360 atgttgtaataatgcccgcttcgatgatcgcttaacaataactcttggttgcaggagta 4419
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25092 ATGTTGAACAATGCAAGGTTTGAAGATCTGCTTTCAACAACATCTGGGCTTCGGCTATA 25151

QY 4420 ggaacttcttagctcaacaagaactcctcttccgaagaattcagttactacagccgc 4479
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25152 GGATCTTTCCTTAGGAAGAAGATATCTGAAATTCGACTCACTACCTTACCTTATCATGGCAGA 25211

QY 4480 ggaacttcagttgccatcgatgccaaacctagacagattttatccttaggagctgcattt 4539
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25212 GGCATACCGCTGCTGTGATGCCAAACCTCGCCCAAGAAATTTATTTAGGACCTGCTTC 25271

QY 4540 agtaagatagttggggaaacccaacccaatcaaaaaatgcatataattacttccataagggc 4599
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25272 AGTCAGGTTTTGGTCAAGCCGAGTCTGAATATCACCTTGACAACTATTAAGCATAAAGGC 25331

QY 4600 tctgagtactcttcaacagcttctctatgagtgaggtataatctctgtattcttctgctcaat 4659
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25332 TCAGTCACTACACAAAGCATCTTTTATGTGCGCAATATCTTCTATTTTCTCCGCGATA 25391

QY 4660 aagcaacatggttgggcacttcttccataacagagctgctgctcctatgacatat 4719
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25392 C-----GGTCTCGGCTATTCTATTCCAAAGGTGTGGCAGCCTATGTTATATG 25439

QY 4720 aaacatgatacaacaacactttaccctctctatccatgaaagaaataaagagatgggaa 4779
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25440 CAACATGACACCAACCACTACTACTCTTCTATTGAAGAAATAATATGCAAACTGGGAT 25499

QY 4780 gatttaggatgttagcgatctctctatctctatcgatcttaagaagaacctctcaaat 4839
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25500 ACATATGCTTGGTATTATGATTCGGTTCAGTGTGGATTTTAAAGAACTCAACCTCAC 25559

QY 4840 tctctaaacggatcacgtctctatgggaactcgagattccagcattccgcagaaaaacag 4899
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25560 TCTACAGCAAGCTTACCTTCTATACAGAGCTGAGTATACCAGATTCGCCAGAGAAA 25619

QY 4900 ttacagaaaatcgatacagatcccaagacacttcgatgattgtgcttacagaaaatctgtcg 4959
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25620 TTACAGAGACTAGACTATGATCTAGATCTTCTCTGCTGCTCTTATGGAAACTTAGCA 25679

QY 4960 ctctctgtggatgcgctgcgaagagctatcgaactgaatattcttattgtataat 5019
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25680 ATTCTACTGGGATCTCTGTAGACGGGATTTAGCTTGGCGTGAGATATTCTATATAAT 25739

QY 5020 aagcttgcttagctacatgcctctctatctacagaaaataacctgctgtaaatatcgg 5079
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25740 AAAGTATCAGCTGCGTACCTCCCTGTGATCTCAGGAATAATCCAAAGACGACCTATGAA 25799

QY 5080 gtattgttcttogaatgaagctggtgaagtattctcgaggtcccaactgaacctctgtct 5139
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25800 GTTCTCTCTACAAAAGAAAAGGCAACGCTAGTCAACGTTCTCCCTACAGAAGAACGCGCT 25859

QY 5140 agagcagaataacagtactcaacataatcttggctcctcttgactctctacagaaactat 5199
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25860 CGTCAGAGGTGAGCTCAAAATTTATTCTTGAAGTACTTGACACTCTTACGGGACGATAT 25919

QY 5200 actatcgatgtaggtatgtatcagctatctcgcaaatgactagctgcggtgctcgatgac 5259
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25920 ACTATTGATGCTTCANTGAATTAATTTAGTGGCAATATGGCCACGGAGGGATCCGGTTTGA 25979
```

```
QY 5260 ttcta 5264
|||||
Db 25980 TTCTA 25984
```

## RESULT 8

```
AC074221 117655 bp DNA HTG 15-NDV-2000
Mus musculus chromosome 15 clone RP23-81A4 strain C57Bl/6J, WORKING
DEFINITION
AC074221
AC074221.5 GI:11177936
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
house mouse.
ORGANISM
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 117655)
Han, J., Montgomery, K.T., Grills, G., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucheralapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 117655)
Han, J., Montgomery, K.T., Grills, G., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucheralapati, R.
Direct Submission
Submitted (20-JUL-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Nov 15, 2000 this sequence version replaced gi:11067106.
-----Genome Center
Center: Albert Einstein College of Medicine
Center Code: AECOM
Web site:
http://sequence.aecom.yu.edu/cgi-
bin/ws.exe/mouseDB/mouseSEQ/mouseseqtable.hts
Contact: jhan@sequence.aecom.yu.edu
-----Summary Statistics
Center project name: ADZ - L08752
Sequencing vector: pUC18; - L08752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 109029 at least Q20
*Consensus quality: 105812 at least Q30
*Consensus quality: 99196 at least Q40
**Estimated insert size: agarose-FP - N/A
**Quality coverage: agarose-FP - N/A
**Quality coverage: 6.2x sum-of-contigs - N/A
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 11535: contig of 11535 bp in length
* 11536: gap of unknown length
* 11556: contig of 8842 bp in length
* 20398: gap of unknown length
* 20417: gap of unknown length
* 28110: contig of 7693 bp in length
* 28130: gap of unknown length
* 28131: contig of 7645 bp in length
* 35775: gap of unknown length
* 35796: contig of 3303 bp in length
* 39099: gap of unknown length
* 39118: contig of 6446 bp in length
* 45564: gap of unknown length
* 45584: gap of unknown length
* 50931: contig of 5347 bp in length
* 50951: gap of unknown length
```









0;

Search completed: October 26, 2001, 22:30:17  
Job time: 13750 sec

	Query Match	1.1%	Score 58.8;	DB 5;	Length 624;
	Best Local Similarity	64.9%	Pred. No. 0.031;		
	Matches 87;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;
Qy	2394	agcaaaactcctccacagaggtcgcaaacacagcttcagcaacctgctgcagctgtgc	2453		
Db	435	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	376		
Qy	2454	ttcctcaagcagcgcgagcagcgccaccatcatctccagcaaacacaaactttatcagg	2513		
Db	375	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	316		
Qy	2514	tgtagtagggaggag	2527		
Db	315	AGTAGTAGTAGTAG	302		

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2001, 08:35:11 ; Search time 18.86 Seconds  
(without alignments)  
1912.741 Million cell updates/sec

Title: US-09-454-684A-180  
Perfect score: 8942  
Sequence: 1 MKWLSATPAVFAVLPSVSGF.....IDVGMVTLQMTSCGARMIF 1752

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	374.5	4.2	1529	2	US-08-728-470-10
2	374.5	4.2	1529	4	US-08-719-641-10
3	373.5	4.2	1536	1	US-08-038-682-2
4	373.5	4.2	1536	1	US-08-302-832-2
5	373.5	4.2	1536	2	US-08-530-198-2
6	373.5	4.2	1536	2	US-08-469-880-2
7	373.5	4.2	1536	2	US-08-728-470-2
8	373.5	4.2	1536	2	US-08-617-697-2
9	373.5	4.2	1536	4	US-08-719-641-2
10	371.5	4.2	1600	2	US-08-617-697-10
11	370.5	4.1	1599	2	US-08-617-697-9
12	366.5	4.1	1338	2	US-08-728-470-9
13	366.5	4.1	1338	4	US-08-719-641-9
14	361.5	4.0	2353	4	US-09-377-155-33
15	361.5	4.0	2353	4	US-08-913-942-4
16	360	4.0	1912	3	US-08-409-995-4
17	360	4.0	1912	3	US-08-465-467-4
18	353.5	4.0	1477	1	US-08-038-682-4
19	353.5	4.0	1477	1	US-08-302-832-4
20	353.5	4.0	1477	2	US-08-530-198-4
21	353.5	4.0	1477	2	US-08-469-880-4
22	353.5	4.0	1477	2	US-08-728-470-4
23	353.5	4.0	1477	2	US-08-617-697-4
24	353.5	4.0	1477	4	US-08-719-641-4
25	352	3.9	1612	1	US-08-169-927-2
26	291	3.3	2123	4	US-08-968-685A-10
27	277	3.1	1026	1	US-08-194-290-7

28	276	3.1	1026	2	US-08-614-377A-7	Sequence 7, Appl
29	276	3.1	1026	4	US-09-142-648B-7	Sequence 7, Appl
30	272	3.0	1160	3	US-08-808-599A-24	Sequence 24, Appl
31	269	3.0	1537	1	US-08-325-267A-2	Sequence 2, Appl
32	260	2.9	1837	3	US-08-928-361B-5	Sequence 5, Appl
33	258.5	2.9	1721	3	US-08-700-651-5	Sequence 5, Appl
34	255.5	2.9	1721	3	US-08-928-361B-6	Sequence 5, Appl
35	254.5	2.8	2035	1	US-08-046-585-5	Sequence 6, Appl
36	254.5	2.8	2035	1	US-08-393-703-5	Sequence 5, Appl
37	254.5	2.8	2035	5	PCT-US93-11721-5	Sequence 5, Appl
38	248.5	2.8	1702	4	US-08-296-791-5	Sequence 5, Appl
39	248.5	2.8	1702	5	PCT-US95-10661A-5	Sequence 5, Appl
40	246.5	2.8	1248	2	US-08-348-353-17	Sequence 17, Appl
41	246.5	2.8	1248	2	US-08-465-965-17	Sequence 17, Appl
42	246.5	2.8	1248	3	US-08-465-966-17	Sequence 17, Appl
43	234	2.6	1848	4	US-08-296-791-6	Sequence 6, Appl
44	234	2.6	1848	5	PCT-US95-10661A-6	Sequence 6, Appl
45	233.5	2.6	1252	2	US-08-682-517-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1  
US-08-728-470-10  
; Sequence 10, Application US/08728470  
; Patent No. 5928651  
; GENERAL INFORMATION:  
; APPLICANT: Barenkamp, Stephen J  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/728,470  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/302,832  
; FILING DATE: 16-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US93/02166  
; FILING DATE: 16-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9205704.1  
; FILING DATE: 16-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berkstresser, Jerry W  
; REGISTRATION NUMBER: 22,651  
; REFERENCE/DOCKET NUMBER: 1038-633  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 415-0813  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1529 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-728-470-10



TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813

TELEFAX: (703) 415-0813

FORMATION FOR SEQ ID NO: 10:

### SEQUENCE CHARACTERISTICS:

**LENGTH:** 1529 amino acids

TYPE: amino acid

**STRANDEDNESS:** single

TOPOLOGY: linear

719-641-10

Query Match 4.28; Score 374.5; DB 4; Length 1529;

Query Match 4.26, Score 374.5, DB 4, Length 1523,  
Best Local Similarity 20.68; Pred. No. 1.9e-16:

Best local similarity 20.00, Recd. no. 1.50 10;  
Matches 308; Conservative 221; Mismatches 597;  
Indels 369; Gaps 66;

```

40  T F T E T I G A G A E Y I V S G N A S P T K T N I P T T D T T P T N S N S S S S G E T A S V S E D S D S T T T T 99
176 T I T Y S I A A P E N E A I N L G D I - F A K G G I - - - N V R A A T I R N K G L S A D - S Y S K D S G N I V L 229
100 P D P K G G A F Y N A H S G V L S F M T R S C T E G S L T I S E I K M T E G G A L F S Q G E L L F T D L T S L I T Q 159
230 S A K E G E A E I - - - G V I S A Q N Q A K G K L M I T G D K V T L T K G A V I - - - - - - - - - - - - - 269
160 N N L S O L S G G A I F - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 194
270 - D L S K E G G E T Y L G D R E G E K N G I Q L A K K T T L E K G S T I N V S K E K G R A I V W G D I A L D 328
195 A P V K P T E P K A O T A S - - E T S C S S S S S S D N S V S P S S R A E P A A N I Q S H F I C A T A P A A Q 252
329 G N I N A Q G S D I A K T G G F V E T S G H D L S I G D D V I V D A K E W L L D P D V S I E T - - L T S G R N T T G 386
253 T D T E T S P S H P G S G G A I Y A K G D L T I A D S O E V L F S I N K A T K D G G A F A E K D V S F E N I T S L 312
387 N O G Y T T G D T E S P K G N S I S K P T L T N S L E Q I L - - - - - - - - - - - - - - - - - - - - - - - 429
313 K V Q T N G A E E K G A I Y A K G D L S I Q S K S O L F N S N Y S K O G G A L Y V E G G I N F Q D L E E I R I K Y 372
430 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 471
373 N K A C T F E T - - K K I T L P S L K A Q A S A G N A D A W A S S P Q S G C A T T V S D G S S S G S - - - D S D 427
472 - K A G S W D V H K N I T L G T G L F N I V A G D S V A F - - - - - E R E G D K A R N A T Q A I T A Q C T I T V N K D 526
428 T S E - - - - - T V P Y T A K G G L - - - - - Y T D K N L S I T N I T G I E I A N K A T D V - - - - - G G G 469
527 D K Q P R F N V S I N G T G K G L K F T A N Q N N T H R F D G E L N I S G I V T I N Q I T K D K Y K W N A S K D S 586
470 A Y V K G T L T C E N S H R L O F L K - - - - - N S S D K Q G - - - - - G G I Y E D N I T L S N T I G K T 513
587 Y M N V S S L T N V P Q R F T F I K F Y D S G S N Q D L R S R R S P A G V H F N G I G K T F N I - G A N A K A 645
514 L F - Q E N T A K E E G G L F I K G T D K A L T W G - - - - - - - - - - - - - - - - - - - - - - - - - - - 549
646 L F K L P N A A T O P K K E L P I T - F N A N I T A T S D G S V M F D I H A N I T S R A A G I N M D S I N I T G G 704
550 - - - - - T S E K H G G G A F V T K E - - - - - I S Q T Y T S - D V E T I P C I T P V H C E T I V T G N 590
705 L D F S I T S H N R N S A F E I K K D L T I N A T O S N F S L O T K D S F N E Y S K H A I N S H N L T I L G C N 764
591 K S T G G - N G G V Y C T K R L A L S N I Q S I S I S G N S A A E N G G A H T C P D S F P T A D T A E O P A A A S A A 649
765 V T L G E N S S S I T G I N I T N K A N Y T L Q - - - - - - - - - - - - - - - - - - - - - - - - - - - 804
650 T S T P K S A P V - - S T A L S T P S S T V S S L T L A A S S O A S P A T S N K E T O P D N A D T D L L I D V Y D 707
805 T L T G N T S V E G N L S T G A N A I V G N L S I A E D S T F K G E A S D N L - - - - - N I T G T F N N G T A N 859
708 T T I S K T A K G G G I Y A K K A M S R I D O L N I S E N A - - - T E I G G I O C K - - - - - E S L E I D 757
860 I N I Q G V V K L O G D I N N K G - - - - - U N I T N A S G T Q K T I I N G I N I N E K G D L N I K K I A D 912
758 A L V S L V S T E N I V L G K E G - G L H A K T V N I S N - - - L K S Q F S N N K A N S S T G V A T T A S A P A A 813

```

REFERENCE/DOCKET NUMBER: 1038-293

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 415-08

TELEFAX: (703) 415-0813

INFORMATION FOR SEQ ID NO:

### SEQUENCE CHARACTERISTICS:

LENGTH: 1536 amino acids

TYPE: amino acid

**STRANDEDNESS:** single

TOPOLOGY: linear

MOLECULE T

Query Match	4.2%;	Score 373.5;	DB 1;	Length 1536;
Best Local Similarity	22.2%;	Pred. NO. 2.2e-16;		
Matches 316;	Conservative 156;	Mismatches 515;	Indels 437;	Gaps 67;

QY	40	TFTTETGEACAEYIVSGNASFFKFTINPTDTHPTPNSNSSSSGSETASVSEDSUSTTTT	99
Dy	248	TITYSTAAPNEAVNLGDI--FAKGNNI---NVRAAATIRNOGLSAD--SVSKDKSGNIYL	301
QY	100	PDPKGGCAFNAHSGYLSWTRSGTSCSLTSBKMTGGGAIF-----S	144
Dy	302	SAREGEAEI----GGVISAQNOAKCGKLMITGDVKVTLTKTGAVIDLSGKEGETYLGDE	357
QY	145	QGE-----LLFDTLSTL---TIQNILSQLSGGAIFOGGTISLSGITKATFCSNSAEPVP	196
Dy	358	RGECKNGIQLAKKTSLEKGTINVSGKEGGRAIVMGDIALDGNINAOGSGDIAKTGGF	417
QY	197	VKPTE-----PQAQTASETSSGSSSGND-----SVSPSSSR	230
Dy	418	VETSGHDLFIKDNAIVDAKEWLLDFDNVSNIAETAGRNSTSEDEYTGSGNSASTPKRK	477
QY	231	AEPAAAAL-----QSHFICATATAAOTDTETS-----TPSHKPGSGGAIYAKGDLT	277
Dy	478	EKTTLNTWLLESILKGTFTVNITANORIYVNSSINLSGLTWBSRSGGVGINNDIT	537
QY	278	IADSQEVFLSINKATXDGGAIFAPEKDVSFE-----NITSKVOTNGABEKGGAIIYAKGLD	332
Dy	538	TGDDTR---GANLIYSGGWVDVHHKNISLGAOQNINITA---KQDIAFEKGSNQVITGO	591
QY	333	SIOSKOSLPEN-SNYSKOGGALYVEGGINFODLEEIRIKYNKAGTFE-----TKKITLP	386
Dy	592	TIISGNOKGRFNFWLUNGTS-----GLQF---TTKRTNKYAITNKFEGTLNISGKNVIS	644
QY	387	SLKAQASAG-----NADAWASSPSQSGATTVSDSGD-----SSSGSDSDTSETVPVTA	436
Dy	645	MVLPNKESGYDKFKGRTYWNLT-----LVNSESGEFNLTIDSRGSDSAGTLTQPYNL	697
QY	437	KGGLYTDK-----NLSITNITGIETAN-----NKATDVGGGAYYKGLTTCBN	480
Dy	698	NGISFNKDDTTFNVARNARVNFIDKAPIGNKYSSLNAYAFNGNISVSGGGSVDFTLLASS	757
QY	481	SH-----RLQFLKNSDDKOGGIYGEDNITLLS-----NLT-----	510
Dy	758	SNVQTPCVWLINSKYFNVTSGSSLRFPKTSGTKGPSI-EKDLTLNATCGNITLLOVEGT	815
QY	511	-----GKTLFOENTAKEBGGGLFIKGTDKALT-MTGUDSFCLINNTSEKHGGGAFVTKEIS	565
Dy	816	DGMIGKIVAKKNIITEPGENITF-GSRKAVTEIEG--NVTINNNAVTLIGSDEFDHQHP	872
QY	566	QTYTSDVETIPGTPVHGERTVIITGNKSTGGNGGVCCTKLALSINLOSISISGNSRAE---	622
Dy	873	LTIKKDY-----IINSIGNITAGCN-----IVNIAGNLTVESNA	905
QY	623	-----NGGHAHTCPDSFFTADTAQPAASAASATSTPKSAPYSTALSTPSSSTVSS	672
Dy	906	NFKAITNFTENVGL-----FDNKGNSNISIAKGCAREFKDINDSKNLSITTNSSSTYR-	958
QY	673	LTLAASSQASPATSNETODPNADTDLLDIYVVDTTISKNTAKGGGIYAKKAKMSRID	732
Dy	959	-----TIISGNITTNKNG-----	970



; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-302-832-2

Query Match 4.2%; Score 373.5; DB 1; Length 1536;
Best Local Similarity 22.2%; Pred. No. 2.2e-16;
Matches 316; Conservative 156; Mismatches 515; Indels 437; Gaps 67;

QY 40 TETETICEAGAEYIVGNSAFTFTNPTDPTPTNSNSSSSGETASVSDSDSTTTT 99
DB 248 TITYSIAAPENEAIVNGDI-FAKGGNI---NVRAAITRNQKLSAD--SVSKDKSGNIVL 301
QY 100 PDPKGGGAFYNAHSGVLSFMSTRGTEGSLTSLBKMTGEGGAIF-----S 144
DB 302 SAKGEAEI---GGVISAQNAQKAGKLMITGDKVTLKTGAVIDLSGKEGGTYLGDE 357
QY 145 QGE-----LLETDTLSL---TIONNLSQLSGAIFGSGTSLSGITKATPSCNSAEVPAP 196
DB 358 RGEGKNGIQIALLKTSLEKGSTINVSKEKGRAIVMGDIALIDGNINAQSGDIAKTGGF 417
QY 197 VKKPT-----PKAOTASETSGSSSSGND-----SVSSPSSSR 230
DB 418 VETSGHDLFTKDAIYDAKEWLLDFDNVSIETAETAGTSNTEDEDTYTGNSASATPKRNN 477
QY 231 AEPAAANL-----QSHFCATATAPAAQPTDTS-----TPSHKPGSGGAIYAKGDLT 277
DB 478 EKTTLNTTLESILKKTGFVNITANQRIYVNSSINLSGSLTLWSEGRSGGVEINNDIT 537
QY 278 IADSQEVLFISINKATDGGGAIFAEKDVFE-----NITSLKVQTNGAEEKGGAIIYAKGDL 332
DB 538 TGGDTR---GANLTIYSGGVWVHKNISLGAQGNINITA---KQDIAFEKGSNOQVITGQ 591
QY 333 SIQSSKOSLPN-SNYSKOGGALYVEGGINFQDLEETRIKYNKAGTFE-----TKKITLP 386
DB 592 TITSGNQKGRFRNVLNLTGTS-----GLQF--TTKRTNKYATINPKFGLNISGKNVIS 644
QY 387 SLKAQASAG-----NADAWASSPQSGGATTVSDSGD-----SSSGSDSDTSETVPVTA 436
DB 645 MVLPKNESGYDKFKGRTYNLTLS-----LNVSEGEFNLITDSRGSAGILTQPYNL 697
QY 437 KGGGLYTDK-----NLSTNTITGIEIAN-----NKATDVGGGAYVKGTLTCEN 480
DB 698 NGISFNKDDTFNVERNARVNFDAKIPIGINKYSLNAYSPFNLSVSGGSGVDFTLIASS 757
QY 481 SH-----RLQELKNSSDKOGGGIYGEDNLTLS-----NLK----- 510
DB 758 SNVQTPGVVINSKYFNVSTGSSSRFTSGSTKGTGFSI--EKDULNATGNTITLLQVEGT 815
QY 511 -----GKTLFOENTAKEEGGLFIKGTDKALT-MTGLDSDFLCINNTSEKHGGGAFVTKEIS 565
DB 816 DGMIGKGIYAKKNITTEGGNITF-GSRKAVTEIEG--NVINNANVTLLIGSDFDNHQKP 872
QY 566 QTVTSVETIPGTPVHGETVITGKNTSGTGNGGVCVTKRLALNSLOISISTGNSAAB--- 622
DB 873 LTIKKDV-----IINSGNLTAGN-----IVNIAGNLTVESNA 905
QY 623 -----NGGGAHTCPDSPFTADAEQPAASAATSTPKSAPYSTALSTPSSSTVSS 672
DB 906 NFKAITNFTFNVGGL-----FONKGNISNISIAGKGAREFKDIDNSKNLSTTNSSTYR- 958

QY 673 LTLAASQASPATSNKETQDPNADTLLIDYVVVDTTISKNTAKKGGIYAKKAKMSRID 732
DB 959 -----TIISGNITNKG----- 970
QY 733 QLNI-SENSATE--IGGGICCKESLELDALVSLVSTENL---VGKEGGGLHAKTVNISNL 786
DB 971 DLNITNEGSDTEMOIGDVSQKEGNLTISSDKINITKQITKAGVDGSDNSDATNNANL 1030
QY 787 K-----SGFSFSNNKA-----NSSSTGVATTASAPAAAAAASLQAAAAAAPS 827
DB 1031 TIKTKELKLTQDLNISGF---NKAETAKDGSDLTIGNTNSADGTNAKV----- 1077
QY 828 SPATPTYSGVVGGAIIY--GEKVTFPQSGTCQFSGNQAIDNNPSSQSLNVQGGAIYAKTS 885
DB 1078 -----IFNQVKDSKISADGHKVTL---HSKVEISGS---NNNTEDSSDNNAGLTIDAKN- 1125
QY 886 LSISSDAGTSYIFSGNSVSTGK-----SQTGGIAGGAIYSPVTLNCPATFNSNTASI- 940
DB 1126 -----VTVNNNITSHKAVSISATSGEI-----TTKTGTTINATTGNVE 1163
QY 941 ATPKTS-----EDGSSGNSIKDTIGGAIACTAITLSCVSRFSFGNTADLGAAGTFLAYANT 996
DB 1164 ITAQTGILGGIESSGSVTLTATEGALA-----VSNISGNT-----VVTANSQA 1209
QY 997 PSATSGSONSITEKITLENGSFIFERQANKRGAIIYSPSVSIRKGNITFNQNTSTHDGSA 1056
DB 1210 LTTLAGSTIKGTSVTTSSOS-----GDIGGTISGTVVEVKA---TESLTTQJNSK 1257
QY 1057 IYTKDKATISLSGLVLTGN-NVTATQAS--SATSGQNTN-TANYGAAIFGDPGTTVUSSQ 1112
DB 1258 IKAT-----TGEANVTSATGTIGGTISGTVNVTANAGDLTVGN--GABINA 1302
QY 1113 TDAILTLASSGNITFNSNLSLONNOGDTPAKPCSIAGYVKLSLOAAKGTISFFDCVHT 1172
DB 1303 TEGAATLTTSSGKLT-----TEASHITSAGQVNLQAQ--DGSVAGSINIANV 1349
QY 1173 STKTKSTQNVYETLIDINKENSNPYTGTIVFSELHENKSYIPQNAIHLNGTLVLK:KKT 1232
DB 1350 TLNWTG-----TLTVKGSNINATSGTLVINA-----KDAELNGAALGNHT----- 1390
QY 1233 ELHVVSFEQKEGKLINEPGAVL-----SNQNIANGALAINGLTI 1272
DB 1391 ---VVNATNANGS-----GSVIATTSRVNITGDLITINGLNI 1425

RESULT 5
US-08-530-198-2
; Sequence 2, Application US/08530198
; Patent No. 5869065
; GENERAL INFORMATION:
; APPLICANT: BARENKAMP, STEPHEN J
; APPLICANT: ST. GEME III, JOSEPH W
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS
; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,198
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: BERKSTRESSER, JERRY W  
REGISTRATION NUMBER: 22,651  
REFERENCE/DOCKET NUMBER: JWB-1186  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1536 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-530-198-2

Query Match 4.2%; Score 373.5; DB 2; Length 1536;  
Best Local Similarity 22.2%; Pred. No. 2.2e-16;  
Matches 316; Conservative 156; Mismatches 515; Indels 437; Gaps 67;

QY 40 TTEIIGEGAGAEYIVSGNASFTKFNIPITDTTPTNSSSSSGSETASVSDSDSTTTT 99  
DB 248 TTYTIAAPENAVNLGDI-FAKGGNI---NVRAATIRNOGKLSAD--SVSKDKSGNIVL 301  
QY 100 PDPKGGAFYNAAHSGVLSFMTSGTEGSLTSEIKWTGEGGAI- - - - -S 144  
DB 302 SAKEGEAEI---GGVISQNOQAGGKLMITGDKVTLTGAVIDLSCGEGETVYLGDE 357  
QY 145 QGE---LLFTDLTSL-----TIONNLSQSGAIFGSGTISLSGITATFCNSAEVAP 196  
DB 358 RGEKNGIQLAKKTSLEKSTINVSKEKGRVAVWGDIALIDGNINAGSGDIAKTGCF 417  
QY 197 VKKPE-----PKQATSETSGSSSSGND-----SVSPSSSR 230  
DB 418 VETSGHDLFIKDAIVDAKEWILDFDNVNSINAEATAGRSNTSEDDYTGSGNSASTPKRKN 477  
QY 231 APAAANL-----QSHFICATAPAAQTDTEIS-----TPSHKPGSGGAIYAKGDLT 277  
DB 478 EKTTLTNTLESILKKGTFVNITANQRIYVNSINLSGSLTLWSEGRSGGVEINNDIT 537  
QY 278 IADSOEVLFSINKAKDGAFAEKDVSE-----NITSLKVOTNGAEKGGAIYAKGDL 332  
DB 538 TGDTR---GANLTIYSGWVDVHKNISLCAOQNITA---KQDIAPEKGSQVITGOG 591  
QY 333 STOSSQSIFN-SNYSKQGGALYVEGGINFODLEIRIKYNKAGTFE-----TKKILP 386  
DB 592 TITSGNQKGFNNVSLNGTGS-----GLQF--TTKRTNKYAITNKFECTLNISGVNIS 644  
QY 387 SLKAQASAG-----NADAWASSSQSGGATTVDSDGD-----SSSGSDSDTSETVPVTA 436  
DB 645 MVLPMNESYDKFGKTYWNLTS-----LNVSEGENLTIDSGSDSAGTLTQPYNL 697  
QY 437 KGGLYTDK-----NLSTINITGIEIAN-----NKATDVGGGAYVKGTLTCEN 480  
DB 698 NGISEFNKOTTFENVERNARVNFIDKAPIGIKYSSLYASFNISVSGGSDVDFTLASS 757  
QY 481 SH-----RLQFLKNSDKQGGIYGEDNITLS-----NLT----- 510  
DB 758 SNVQTPGVVINSKYFNVSTGSSLRFTKSTGKTFISI--EKDLTLNATGNITLLQVET 815  
QY 511 ----GKTLFOENTAKEGGGLIKGTDKALT--WTGLDPSFLINNTSEKHGGGAFVTKETS 565  
DB 816 DGMICKGIVAKKNITFEGGNITF-GSRKAVTEIEG--NVTINNANVTILIGSDFDNHQP 872  
QY 566 QYVTSVETIPGIVPHGVETVITGNKSTGGGCVCTKRLALSNLOSTISGNSAAE--- 622  
DB 873 LTIKKDV-----IINSGLNLTAGN-----IVNIAGNLAVESNA 905  
QY 623 -----NGGGAHTCPDSEPTADTAQPAASAASTPKSPAVSTALSTPSSSTVSS 672  
DB 906 NPKATNTFTFNVGGL-----FDNKGNSNISIAKGAERPKDIDNSKNLSITTSSTVTR- 958

QY 673 LTLAASSQSPATSNKETQDPNADTDLILDYVVDTTISKNKAKKGGIYAKKAKMSRID 732  
DB 959 -----TILSGNITNKG----- 970  
QY 733 QLNI--SENSATE--TGGICCKESLELDALVSLSVTEML---VGKEGGGLHAKTVNISML 786  
DB 971 DLNITNEGSDTEMOIGGDVSOKEGNLTISDDKINTTKQITIKAGVDGENSDSDATNNANL 1030  
QY 787 K-----SGFSFSNNKA-----NSSSTGVATTASAPAAAAASLOAAAAAPS 827  
DB 1031 TIKTKELKLTQDLNISGF-----NKAEITAKGDSDLTIGNTSADGTNAKKV----- 1077  
QY 828 SPATPTYGVGGAIY--GEKVFSCQCTCFSGNOAIDNNPSSQSLNVQGGAIYAKTS 885  
DB 1078 -----TFNOVKDSKISADGHKVTL---HSKVETSGS---NNNTEDSSONNAGLTIDAKN- 1125  
QY 886 LSIGSDAGTSYIFSGNSVSTGK-----SOTTQIAGGAIYSPVTILNCPATFSNWTASI- 940  
DB 1126 -----VTNNNITSHKAVSISATSGEI-----TTKTGTTINATTGNVE 1163  
QY 941 ATPKTS-----EDGSSGNSIKDTIGGATAGTAITLSGVSRFSGNTADLCAAIGTLANANT 996  
DB 1164 ITAQTGSLIGLIESSGSVTLTATEGALA-----VSNISGNT-----VTVTANSGA 1209  
QY 997 PSATGSGNSITEKITLENGSFIFERNOANKRGAIYSPSVSIKGNITFNQNTSTHDGSA 1056  
DB 1210 LTTLAGSIKGTESVTTSSQS-----GDIGTISGGTVEVKA-----TESLTTQSNK 1257  
QY 1057 IYFTKDATIESLGVLFNGN-NVTAQAS--SATSGQNTN-TANYGAAIFGDPGTTQSSQ 1112  
DB 1258 IKAT-----TCEANVTSATIGTIGTSNTVNTANAGDLTVGN--GAEINA 1302  
QY 1113 TDAIITLLASSNITFSSNLSLONNOGDTSPAKFCSIAGYVVKLSLOAAKGTISFFDCVHT 1172  
DB 1303 TEGAATLTSSKLT-----TEASSHITSAGQOVNLSAO--DGSVAGSINAANV 1349  
QY 1173 STFKTGSTQNVYETDINKENSNPYTGTFVPSSELHENKSYIPONAILHNGTLVLKKEKT 1232  
DB 1350 TLNWTG-----TLTTVKGSINATSGTLVINA-----KDAELNGAALGNHT----- 1390  
QY 1233 ELHVVSFOKEGSKLIMEPGAVL-----SNQNTANGALAINGLTI 1272  
DB 1391 ---VYNATNANGS-----GSVIATTSRRVNTIGDLITINGLNI 1425

## RESULT 6

US-08-469-880-2  
; Sequence 2, Application US/08469880  
; Patent No. 5876733  
; GENERAL INFORMATION:  
; APPLICANT: Barenkamp, Stephen J.  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; TITLE OF INVENTION: of No. 5876733-Typeable Haemophilus  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; STREET: Bldg. 1  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,880  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9205704.1

;; FILING DATE: 16-MAR-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US PCT/US93/02166  
;; FILING DATE: 16-MAR-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/302,832  
;; FILING DATE: 16-SEP-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Berkstresser, Jerry W  
;; REGISTRATION NUMBER: 22,651  
;; REFERENCE/DOCKET NUMBER: 1038-516 MIS:vg  
;; TELEPHONE: (703) 415-0810  
;; TELEFAX: (703) 415-0813  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1536 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-08-469-880-2

Query Match 4.2%; Score 373.5; DB 2; Length 1536;  
Best Local Similarity 22.2%; Pred. No. 2.2e-16;  
Matches 316; Conservative 156; Mismatches 515; Indels 437; Gaps 67;  
QY 40 TPTETIGEAGAEYIVGNSAFKFTNIPITDTTPTNSNSSSSGSETASVSDSDSTTTT 99  
DB 248 TITYTAAPENEAUNLGDII-FAKGGNI---NVRATIRNOGLSAD--SVSKDKSGNIVL 301  
QY 100 POPKGGGAFYAHSGVLSFWTSGTSLSEIKMTGEGGAIFF-----S 144  
DB 302 SAKGEAEI---GGVISAQNOQAKGKLMITGDKVTLTGAVIDLSGKEGGETYLGDE 357  
QY 145 QGB-----LFTDLTSL-----TTONNLSQLSGGAIFGSGTISLSGITKATFSCNSAEVPAP 196  
DB 358 RGEKNGIQLAKKTSLEKSGTINVSKEKGRGIVMGDIALIDGNINAOOGSGDIATGGF 417  
QY 197 VKKPTB-----PKAQTASSETSGSSSSGND-----SVSPSSSR 230  
DB 418 VETSGHDLFIKDAIYDAKEWLLDFDNVINAETAGRSNTSEDEYTGSGNSASTPRNK 477  
QY 231 AEPAAANL-----OSHFICATAPAAQTDTS-----TPSHKPGSGGAIYAKGDLT 277  
DB 478 EKTTLTNTTLESILKGTFTVNITANQRIYVNSSINLSGLTLWSESGSGGVEINNDIT 537  
QY 278 IADSQEVLFSINKATDKGGAIFAEDVSE-----NITSLKVQTNAGABKGGAIYAKGDL 332  
DB 538 TGDGTR---GANLTIYSGGWDVHKHNSISLGAQGNINITA---KQDIAFEKGSNOQVITGOG 591  
QY 333 SIQSSKOSLFN-SNYSKQGGALYVEGGINFQDLERIKYNKAGTFF-----TKKITLP 386  
DB 592 TITSGNQKGRFNNSLNTGTS-----GLQF---TTARTNKYAITNKEFTLNISGRVNIS 644  
QY 387 SLKAQASAG-----NADAWASSPQSGGATTVSDSGD-----SSSGSDSDTSETVPVTA 436  
DB 645 MVLPKNESGYDKFKGTYNWLS-----LNVSSEGEFNLTIDSRGSDSAGTLTQPYNL 697  
QY 437 KGGGLYTDK-----NLSTNTTIGIEIAN-----NKATDVGGGAYKVGTLTCEN 480  
DB 698 NGISFNKDDTFNVERNARVNFIDKIPAGINKYSSLNAYAFNGNISVSGGSDVFTLLASS 757  
QY 481 SH-----RLOFLKNSDKOGGGIYEDNITLS-----NLT----- 510  
DB 758 SNVQTPGVVINSKYFNVSSTLRFKTSSTGTGFSI---EKDLTLNATGNITLLQVEGT 815  
QY 511 ----GKTLFOENTAKEEGGLFKGTGDKALT-WTGLDLSFCLINNTSEKHHGGAFVTKETS 565  
DB 816 DGMICKGIYAKKNITPEGGNITF-GSRKAVTEIEG---NVTINNANVTILGDFDHNQRP 872  
QY 566 QTYTSDVETIPGITPVHGETVITGNKSTGGNGGVCTKRLALSNLOSISISGNSAAE--- 622

DB 873 LTIKDV-----IISGNLTAGN-----IVNIAGNLTVESNA 905  
QY 623 -----NGGGAHTCPDSPPTADTAQPAASAATSTPKSAPYSTALSTSPSSSTFVSS 672  
DB 906 NFKAITNFTFNVGGL-----FDNKGNISISIAKGGARFKDIDNSNLSITTTSSSTYR- 958  
QY 673 LTLAASSQASPATSNKETQDPNADTDLIDYVVDTTTISKNTAKKGGIYAKKAKM3RID 732  
DB 959 -----FIISGNITKNG----- 970  
QY 733 QLNI-SENSATE---ICGGICCKRESLELDALVLSVTENL---VGKEGGGLHAKTVNLSNL 786  
DB 971 DLNITNEGSDTEMQIGGDVDSQKEGNLTISSDKINITKQITIKAGVDGSDSDATNNAL 1030  
QY 787 K-----SGFSFNKA-----NSSSTGVATTASAPAAAAAASLQAAAAAPS 827  
DB 1031 TIKTKELKLTQDLNISGF-----NKAEITAKDGDSDLTIGVNTNSADGTNAKV----- 1077  
QY 828 SPATPYTSGVVGAIY--GEKVTFSCSGTCQFSGNQAIIDNNPSSQSLNVQGGAIYAKTS 885  
DB 1078 -----TFNOVKDSKISADGHKVTL---HSKVETSGS---NNNTEDSSDNAGLTIDA-KN- 1125  
QY 886 LSISSDAGTSYIFSGNSVSTGK-----SQTGOIAGGAIYSPVTLNCPATFSNNTASI- 940  
DB 1126 -----VTVNNITSHKAVISATSGEI-----TTKTGTGINATTGNVE 1163  
QY 941 AYPKTS-----EDGSSGNSIKDIFGAIAGTATLTSVSRFSGNTADLGAAGTLANANT 996  
DB 1164 ITAQTGSIILGTESSSGSVTLRATEGALA-----VSNISGNT-----VTVTANSGA 1209  
QY 997 PSATGSONSITEKITLENGSFIFERNOKRGAIYSPSVSIKGNITNONTSTHGS 1056  
DB 1210 LITLAGSTIKGTESVTTSSQS-----GDIGGTISGGIVEVKA-----TESLTTQNSK 1257  
QY 1057 IFTKDATIESLGSVLFTGN-NVTATQAS--SATSQONTN-TANYGAAIFGDPGTTQSSQ 1112  
DB 1258 IKAT-----TGEANVTATGTTIGTISGNTVNTANAGDLTVGN--GAEINA 1302  
QY 1113 TDAITLLASSGNITFSSNLSQNOGDTTPASKFCSTAGYVKLSLOAAKGTISFFDCVHT 1172  
DB 1303 TEGAATLTITSSGKLT-----TEASSHTSAKQVNLQAQ--DGSVAGSINAANV 1349  
QY 1173 STKKTGSTQNVYETLDINKENSNPYGTIVFSELHENKSYIPQNAIHLNGTLVLKEKT 1232  
DB 1350 TLTNTTG-----TLTVKGSINATSGTLVINA-----KDAELNGAALGNHT----- 1390  
QY 1233 ELHVVSFEQKEGSKLIMEPAVL-----SNQNTANGALAINGLTI 1272  
DB 1391 ---VWATNANGS-----GSVIATTSSRVNITGDLITINGLNI 1425

RESULT 7  
US-08-728-470-2  
; Sequence 2, Application US/08728470  
; Patent No. 5928651  
; GENERAL INFORMATION:  
; APPLICANT: Barenkamp, Stephen J  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; TITLE OF INVENTION: of No. 5928651-typeable Haemophilus  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/617,697  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,832  
FILING DATE: 05-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US93/02166  
FILING DATE: 16-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Berkstresser, Jerry W  
REGISTRATION NUMBER: 22,651  
REFERENCE/DOCKET NUMBER: 1038-557  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1536 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-617-697-2

Query Match 4.2%; Score 373.5; DB 2; Length 1536;  
Best Local Similarity 22.2%; Pred. No. 2.2e-16;  
Matches 316; Conservative 156; Mismatches 515; Indels 437; Gaps 67;

QY 40 TPTETIGEAAGVIVSGNASFTFTNPTPTDTPNNSSSSSGETASVEDSDSTTT 99  
DB 248 TITYSTAAPNEAVNLGDI-FAKGNI--NVRAATIRNOCKLSAD--SVSKDKSGNVL 301  
QY 100 PDPKGGAFYNAHGSVLFMRSGTGLSLPLBIKMTGEGGAIF-----S 144  
DB 302 SAKGEAEI---GGVISAQOAKGKLMITGDKVTLTKGAVIDLSGKEGETYLGDE 357  
QY 145 QGE---LLFTDLTSL---TIONNLSQLSGGAIFGGSTISLSGIYKATFSCNSAEVPAP 196  
DB 358 RGEKNGIQIQAATKLEKGTINVSKEKGGAIVWGDIALIDGNINAQSGGDIATGGF 417  
QY 197 VKKPT-----PKAOPASSETSGSSSSGND-----SVSSPSSR 230  
DB 418 VETSGHDLFIKDNAIVDAKEWLLDFDNVSNATAGRSNTSEDEYTGSGNSASTPKRNK 477  
QY 231 AEPAAANL-----QSHFTCATAPAAOPDTETS-----TPSHKPGSGGAIYAKGDLT 277  
DB 478 EKTTLNTTLESILKKGTFVNITANQRIYVNSSINLSGSLTLWSEGRSGGVEINNDIT 537  
QY 278 IADSQEVLFSINKATRDGGAIFAEDKVSFE-----NITSLKVQTNAGAEKGGAIYAKGDL 332  
DB 538 TGDQTR---GANLTIYSGGWVDVHKNIISLGAQGNINITA---KQDIAFERGSNOQVITQG 591  
QY 333 SIOSSKQSLFN-SNYSKQGGALYVEGGINFQDLEETRIKYNKAGTPE-----TKKITLP 386  
DB 592 TITSGNQKGERFNNVSLNGTGS-----GLOF--TTRTKNYATNKFEGTLNLSGRKNIS 644  
QY 387 SLKAQASAG-----NADAWASSPQSGGATTVSDGD-----SSGSDSDTSETVPVTA 436  
DB 645 MVLPKNESGVDFKPGRTYNWLTGS-----LNVSSEGFNLITDSRGSAGTLTPYNL 697  
QY 437 KGGGLYTDK-----NLSTNTITGIEIAN-----NKATDVGGGAYVKGITTCEN 480

Db 698 NGISFNKDDTTFNVARNVDFDIKAPIGINKYSSLANVAFNGNISVSGGSDVDFTLASS 757  
QY 481 SH-----RLQFLKNSDKOGGGIYGEDNITLS-----NLT----- 510  
Db 758 SNVQTPGVINSKYFNVSTGSSLRFKTSGTKGFSI--EKDLTLNATGNTITLQVEGT 815  
QY 511 -----GKTLFOENTAKREGGLFIKGTDKALT-MTGLDLDFCLINNTSPKHGGGAFTVKEIS 565  
Db 816 DGMIGKGIYAKKNITPEGGNITF-GSRKAVTEIEG--NVTINNANVTLLIGSDFDNHQKP 872  
QY 566 QTYTSDVETIPGTPVHGETVITCNKSTGGNGGVCYKRLALSNIQISISGNSAAE--- 622  
Db 873 LTIKKOV-----IINSGNLTAGN-----IVNTAGNITVESNA 905  
QY 623 -----NGGAHTCPDPSFTADTAEPAAASAATSTPKSAPVSTALSTPSSSVSS 672  
Db 906 NFKAITNFTNVGGL-----FDNKGNSNISIAKGGARFKDIDNSKNLSITNSSSYR- 958  
QY 673 LTLAASSQASPATSNKETQDPNADTDLIDYVVDVTTISKNTAKKGGIYAKKAKMRID 732  
Db 959 -----TIISGNITNKG----- 970  
QY 733 QLANI-SENSATE-IGGICCKESLELDALVLSVTENL---VCKEGGGLHAKTVNISNL 786  
Db 971 DLNITNEGSDTEMOIGDVSQKEGNLTISDKINITKQITIKAGVDGENSDSDATNANL 1030  
QY 787 K-----SGFSFSNNKA-----NSSSTGVATTASAPAAAAASLQAAAAAAPS 827  
Db 1031 TIKTKELKLTQDLNISGF---NKAETAKDGSDLTIGTNTNSADGTNAKV- 1077  
QY 828 SPATPYTGVVGGAIY--GEKVTFQSCGFCQFSGNOAIDNNPQSSSLANVOGGGAIYAKTS 885  
Db 1078 -----TENQVKDSKISADGHKVTL---HSKVETSGS---NNNTEDSSDNNAGLTIDAKN- 1125  
QY 886 LSTGSSDAGTSYIFSGNSVSTGK-----SQTGQTAGGAIYSPVTLNCPATFSNNTASI- 940  
Db 1126 -----VTVNNHNTSHRAVSIATSGEI-----TTKTGTINATTGNVE 1163  
QY 941 ATPKTSS-----EDGSSGNSIKDTIGGAIAGTAITLGSVRSFGNTADLGAAGTLANANT 996  
Db 1164 ITAQTGILGIGESSGSVLTATEGALA-----VSNISGNT-----VVTANSGA 1209  
QY 997 PSATSGSONSITEKITLENGSIFIPERNOAKRGAIYSPSVSIKGNNTTFNQMTSTHGS 1056  
Db 1210 LTLTAGSTIKGTESVTTSSQS-----GDIGGTISGTTVEVKA-----TESLTTQNSK 1257  
QY 1057 IYFTKDATIESLSGVLEFGN-NVTATQAS--SATSGQNTN-TANYGAAITGDPDPTQSSQ 1112  
Db 1258 IKAT-----TGEANVTISATGTIGGTISGNTVNTANAGDLTVGN--GAEINA 1302  
QY 1113 TDAILTLASSGNITFNSNLSQNNQGDTPASKFCSIAGYVKLSLQAAKGTISFFDCVHT 1172  
Db 1303 TEGATLTITSSGKLT-----TEASSHITSAGOVNLSAQ--DGSVAGSINANV 1349  
QY 1173 STKKTGTQNVYETLDINKEENSNNPYTGTIVFSELHENKSYIPQNAIHLNGTLVLKEKT 1232  
Db 1350 TLNNTG-----TLTVKGSINATSGTLVINA-----KDAELNGAALGNHT----- 1390  
QY 1233 ELHWVSFEQKEGSKLINEPGAVL-----SNQNIANGALAINGLTI 1272  
Db 1391 ---VVNATNANGS-----GSVIATTSRRVNTGDLTITINGLNI 1425

RESULT 9  
US-08-719-641-2  
; Sequence 2, Application US/08719641  
; Patent No. 6218141  
; GENERAL INFORMATION:  
; APPLICANT: Barenkamp, Stephen J  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; TITLE OF INVENTION: of No. 6218141-Typeable Haemophilus  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Shoemaker and Mattare, Ltd.  
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
STREET: Bldg. 1  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/719,641

Query Match	4.2%;	Score 373.5;	DB 4;	Length 1536;
Best Local Similarity	22.2%;	Pred. NO. 2.2e-16;		
Matches 316;	Conservative 156;	Mismatches 515;	Indels 437;	Gaps 67;

[illegible]

Patent No. 5977336  
GENERAL INFORMATION:  
APPLICANT: Barenkamp, Stephen J  
TITLE OF INVENTION: High Molecular Weight Surface Proteins  
TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Shoemaker and Mattare, Ltd.  
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
Bldg. 1  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/617,697  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,832  
FILING DATE: 05-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US93/02166  
FILING DATE: 16-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Berkstresser, Jerry W  
REGISTRATION NUMBER: 22,651  
REFERENCE/DOCKET NUMBER: 1038-557  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1600 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-617-697-10

Query Match 4.2%; Score 371.5; DB 2; Length 1600;  
Best Local Similarity 20.7%; Pred. No. 3.2e-16;  
Matches 310; Conservative 220; Mismatches 591; Indels 377; Gaps 67;  
QY 40 TPTETIGAEYIVSGNASFTKFTNPTDTPPTNSNSSSSGETASVSDSDSTTTT 99  
DB 248 TITYSTAAPENEAINGDI-FAKGNI---NVRAATIRNKGKLSAD--SVSKDKSGNVL 301  
QY 100 PDPKGGAFYNAHGVLSFWRSGTGLSLSEIKMTGEGGAIFSGOGLLFTDLTSLTIQ 159  
DB 302 SAKGEAEI---GGVISAQNOQAKGKLMITGDKVTLTKTAVI----- 341  
QY 160 NNLSQLSGGAIF-----GGSTISLSGITK---ATFSCNSAEPV 194  
DB 342 -DUSGKEGGETYLGGBERGKNGIOLAKKTTLEKSTINVSKEKGGRAIVYGDIALID 400  
QY 195 APYKKPTEPKAQTAS--ETSGSSSGNSDVSVPSSSRAPAAANLQSHFICATATPAAQ 252  
DB 401 GNINAQSDIAKTGGFVEVSGHDLISGDDVIVDAKELWLLDPPDDVSIEI--LTSGRNNTGE 458  
QY 253 TPTETTPSHKPGGGAIAKAGDLTIADSOEVLFSINKATKDGGAIFAEDKVSFENITSL 312  
DB 459 NQGYTTGDTGKESPKGNSISKPLTNTSTLEQIL-----RRGSYVNITA- 501  
QY 313 KVTNCAEEKGGAIAKAGDLISIOSSKQSLFNSNYKGGGALYVEGGINFDLEEIRIKY 372  
DB 502 -----NNRIYVNSINLSGLTL-----HTRDGG--VKINGDITSNENGLUTI-- 543

QY 373 NKAGTFET--KKITLPSLKAQASAGNADAWASSPOSGSGATTVSOGSDSSGS---DSD 427  
DB 544 -KAGSVVDVHKNTITLGTGFLNIVAGDSVAF-----EREGDKARNATDAQITFAQGTITVWKD 598  
QY 428 TSE-----TVPVYAKGGGL-----YTDKNLSITNITGIIETANKKATDV-----GGG 469  
DB 599 DKQFRFNNSVSLNGTGKGLKFIANNONNPTFKPDGEINISGIVTINQTTKDKVKYWNASKDS 658  
QY 470 AYYKGTILTCENSHRLOFLK-----NSSDKQG-----GGIYGEDNITLSNLTKGT 513  
DB 659 YWNVSSLTLNTVQKFTFIKFDVDSGSGNCDLRSRRSFAGVHENGIGGKTFFNI-GANAKA 717  
QY 514 LF--QENTAKEEGGLFIKGTDRKALMTWG-----LDSFCLINN 549  
DB 718 LFKLPNAATDPKKELPIT-FNANITATGNSDSSVMFDIHANLTSTRAAGINMDSINITGG 776  
QY 550 -----TSEKHGGGAFVTK-----ISQYTS--DVETIPGITPVHGETVITGN 590  
DB 777 LDFSITSHNRNSNAFEIKKDLTINATGNSNFSLKQTSFYNEYSKHAINSNHLTILGNG 836  
QY 591 KSTGG-NGGGVCTKRLALSNLQISISGNSAAENGGGAHTCPDSTPPTADTAEQPAAASAA 649  
DB 837 VTLCGNSSSSITGNINITNKANVTIQ-----ADTSNNTGUKKR 876  
QY 650 TSTPKSAPV--STALSTPSSSTVSSLTLLAASSQASPATSNKE---TQDPNADTDLDIDY 704  
DB 877 TLTGNISVEGNLSLTGANANIVGNLSIAEDSTFKGEASDNLNITGFTTNGTANINIKG 936  
QY 705 VDTTISKNTAKKGGIYAKKMSRIDQLNISENSA---TEIGGGICK-----ESL 754  
DB 937 VVKLGDIINN--KGG-----LNTTASGTOQTIIINGNITNEKGDNL-KNI 979  
QY 755 ELDALVSLSVTENLVGKEGG-GLHAKTVNISN--LKSGFSSNNKANSSTGCVATTASA 810  
DB 980 KADA--EIQIGNISOKEGNLTSSDKVNITNQITKAGVEGGRSDSSEANANLTIOTK 1037  
QY 811 PAAAAASLQAA-----AAAPSPSPATPTYGVGVGAIYGEKVTFSQCSG----- 854  
DB 1038 ELKLAGDLNISGFNKAEITAKNGSDLTIGNASGNDAKAKVTFDKVKDKSISTDGHNVTL 1097  
QY 855 -----TCQFSGNOAIDNNP-----SQSLNVQGGAIYAKTSLISGSSDAGTSYIFSGNSV- 904  
DB 1098 NSEKTSNGSSNAGNDNSTGLTISAKDVTYNNVNTSHKT-INI-SAAAGNVTTKEGTIN 1155  
QY 905 -STGKSQTTQGIAGGAIYSPVTNLCNCPATFSNN-----TASIAPTKTSBDESGSN 954  
DB 1156 ATTGSEVETAQ--NGIKGNITSONVTVTATENLVTTENAVINATSGTVNISTKTGKIG 1213  
QY 955 SIKDTIGGATAGTAITLSG-----VSRESGN-----TADLGAAGICTLANANTPSATSGSNS 1006  
DB 1214 GIESTG-----NVNITASGNLTKVSNITGQDVTVTADAGAL--TTTAGSTISATTGNAN- 1266  
QY 1007 ITEKITLENGSFIFERNQANKRGAIYSPSVSIKGNNTFNQNTSHDGSIAIYFKDATIE 1066  
DB 1267 -----ITTKTGIDNGKVBESSGSVTLVATGATLAVGNISGNTVTITADSGKLTSTVGSTIN 1322  
QY 1067 SLGSV-----LFTGN--NVTATQ-----ASSATSGQNTNTANY 1097  
DB 1323 GTNSVTSSSGDIEGTISGNTVNTASTGDLTIGNSAKVEAKNGAATLTAESKLTQT 1382  
QY 1098 GAATFGDPGTQSSQTDAILTLASSGNITFNNLSNQQODTPASKFCSITAGVVKLSLQ 1157  
DB 1383 GSSTSSNGOFTLTAKDSSIAGNINAAVNTLTGTLTTTGD---SKINATSGTLTINAK 1439  
QY 1158 AAK-----GKTSIFDCVHTS-----TKKTGSTQNVYETLDINKKEENSNPYTGTIVFS 1205  
DB 1440 DAKLDGAASGDRIV--VNATNAGSGNVTAKTSSSVNI--TGDLN-----TINGL 1485  
QY 1206 SELHENKSIYPOANILHNGTFLVLEKTELHVVSFEQESKGLIMEPCAVLSN---QNZAN 1262  
DB 1486 NIISEN-----GRNVTURLRKEIDVKYIQPGVASVEEVIEAKRVLEKVKDLSDEERETIAK 1541  
QY 1263 -GALAI-----NGLTIDLSSMGTTPQAGEIFSPPELRIVATTSASGGSGVSSSIPTN 1313

Db 1542 LGVSAREPVEPNNAITVNTQN-----EFTTKPSSQVTISBKGACFSGNGARVCTN 1592

## RESULT 11

US-08-617-697-9

; Sequence 9, Application US/08617697

; Patent No. 5977336

; GENERAL INFORMATION:

; APPLICANT: Barenkamp, Stephen J

; TITLE OF INVENTION: High Molecular Weight Surface Proteins

; TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Shoemaker and Maltare, Ltd.

; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza

; STREET: Bldg. 1

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202-0286

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/617,697

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/302,832

; FILING DATE: 05-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US PCT/US93/02166

; FILING DATE: 16-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Berkstresser, Jerry W

; REGISTRATION NUMBER: 22,651

; REFERENCE/DOCKET NUMBER: 1038-557

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 415-0810

; TELEFAX: (703) 415-0813

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1599 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-617-697-9

Query Match 4.1%; Score 370.5; DB 2; Length 1599;

Best Local Similarity 19.9%; Pred. No. 3.7e-16;

Matches 317; Conservative 227; Mismatches 585; Indels 465; Gaps 69;

QY 30 FSRVETSSSTTETIGEGAGAYIVYSGNASTFKFTNPIPTDPTTPTNSNSSSSGETASV 89

Db 125 FNRVTSQISQLKILDSNGQVFLNPNGI-----TIGDAIINNGFTAS--TLDI 174

QY 90 SEDSDSTT--TPDPKGGAGFYNAHSGVLSFMRSGTEGSLTSEIKMTGEGGAIFSQOE 147

Db 175 SNEIKARNFTLEQTKDRLAEIVNHLITV---GKDGSVNLI-----GKVKNEG 223

QY 148 LLFTDLTSLTIONNLSQLSGGAI--FGSTISLSGITATFSCNPAEPAP----- 196

Db 224 I-----SVNGGSISLLAQKITISDIINPTIYS---IAAPENEAINLGD 265

QY 197 -----VKKPTPEKQATAGTSGSSSSGSDVSSPSSRAE-----PAAANLQSHFI 243

Db 266 IFPAKGNINVRATIRNKKLSADSVSKDKSN-IVLSAKEGEAEIGGVISAQNOQA--- 321

QY 244 CATATPAAQTDTETTPSHKPGSGGAIYAKGDLTIADSOEVLFSINKATKDGGAIFAEKD 303

Db 322 -----KGGKLMITGDKVTLTKGAV---IDLSEKGEGETYLGSD 356

QY 304 -----VSPENITSLK-----VQINGAEKGGAIYAKGDLISQSSKQSLFNSNYSKQG 351

Db 357 ERGEGKNGIQLAKKTTLEKGSTINVSQ--KEGGRAIWGDIALIDGINHNAQKDIKTG 415

QY 352 GALYVEGGINFODLEIRIKYNKAGTFETKITL--PS-----LKAQASAGNADAWASS 404

Db 416 ---FVETSGHYLSIDDNAIVTKEWLLDPENVTIEAPSASRVELGADRSHSAEVIKVT 472

QY 405 POSGSGATTVSDS-----GDSSSGSD 425

Db 473 KKNNTSLATLTNTTISNLLKSAHVVNITARRKLTVNSSISIERGSHLILHSEGGQGGV 532

QY 426 SOTSETPVPTAKGGGL-----YTKNLSITNITGIIETIANKKATDVGGAY----- 471

Db 533 IDKD-----ITSEGNLTIYSGWVDVHKKNITLGSGLNI-----TTKEGDIAPEDKSGRNN 584

QY 472 ----VKGTILTCENSHRLOFLKNSDKQGGIYGEDNITLSNLTGKTLFOE----- 517

Db 585 LTITAGGITSONSGFRE-----NNVLSNLGGKLSFTDSREDRCRRTK 629

QY 518 -NTAKEEGGLFIKGT-----DKALTMTGLDGDFCLINNTSEKHGGAGFTV 561

Db 630 GNISNKFDTGLNISGTVDISMKAPKVSWFYRDKRTYWNVT---LNVTS-----GSKF-- 680

QY 562 KEISQTYTSDVETIPGI--TPVHGETVITGNKSTGCGGCVCTKRLALNSLOISISGNS 619

Db 681 -NLSIDSTGSGSTGPSIRNAELNG---ITFNKATNIAQGSTANFSAIKAMPFKNANY 736

QY 620 AAE-----GGGHTCPDSTADTAEPAAAAAATSTPKSAPVSTALSTPSSSTVSSL 673

Db 737 ALFNEDISVSGGS-----VNFKNASSNIQTPGVIIKSONFVSGGSTLN-- 783

QY 674 TLLAASSQASPATSNKETQDPNADTDLIDYVDTTIS-----KNTAKGGGIY-- 722

Db 784 --LKAEGSTETAFSTENDLNLNATGNITIROVEGDSRVNGVAAKNITFKGGINTEG 841

QY 723 AKKAKMSRIDQLNISEN-----SATEIGGGICCKESLELDALVSLVTENL 768

Db 842 SOKATTEIKGNVTINKNTNATLRGANFAENKSPLNAGNVINNGNLTTAGSI--INIA 900

QY 769 VGKEGGGLHAKTVNISNLKSGFSFNKA-----NSSSTGVATTASAPA 812

Db 901 TVSKGANLQAITNYTFNVAG--SFDNNGASNISIRAGGAKFKPDINNTSLNITTTSDT-- 956

QY 813 AAAASLQAAAAAPSPATPYSGVVGGAIFYGEK-----VTFSQCSGT 855

Db 957 -----TYRTIIKGNISNKGDLNIIDKKSDAEIIGCNISQKEGN 996

QY 856 COFSGNQAIIDNPNSSSLNVQGG-----AIYAKTSLSIGSSD---AGTSYVIFSGN---SV 904

Db 997 LTISSDKVNITNOITIKAGVEGRSDSSEANANTIQTKELKLAGDLNLSFNKAEITA 1056

QY 905 STGKSQTTQGIAGGAIYSTVTLNCP-----ATFSNNTASIAATPKTSSBDGSGNSIKDT 959

Db 1057 KNGSDLTIGNSGGNADAKKVFDFKVKDSKISTDGHNVTLNSEVKTSGNSAGND--NS 1114

QY 960 IGGAIAGTAITLGSVSRFSNGTADLGAATG--TLANANTPSATSGNSQNSITBKITLENG 1017

Db 1115 TGLTISAKDVTYNN--NVTSHKTNISAAAAGNVTREKGTINATGGS-----VEVTAQNT 1168

QY 1018 FIFERNQAKRAIYSPSVSIKG--NNITPNQNT--STHDGSAIYFTKDIATIESLGSVLE 1074

Db 1169 I-----KGNITSONVTVTATENLVTENAVINATSGVTNISTKTGDIK--GGIEST 1217

QY 1075 GNNVTAT-----QASSATSGONTNTANYGAAIFGDPGTQSSQTDAILTLSSAGNI 1128

Db 1218 SGVNLTASGNTLVKSNITGQDVTYADAGL-----TTTAGST---ISATIGNANIT 1268

QY 1129 SNNSLONNGDTPAKSFCSIAGYVKLSLOAAKGTISFFDCVHTSTKTKGTONVYETLD 1188



Db 1269 KTGDI-NGRVSSGVTIVATCATLAVGNISGNTVTITADSGKLTSTVGSTIN--GTNS 1325  
QY 1189 INKEENSPTGTIVFESSELHENKSYIPONAILHNGTLVLKKEKTELHVVSFFQKEGS--- 1245  
Db 1326 VTTSSQSGDIEGTI-----SGNTVNTASTGDLTIGNSAKV-----EAKNGAATL 1370  
QY 1246 -----KLIMEPGAVLSNQ-----IANGALA--INGLIDISSMGTPOAGEIFSPPELR 1292  
Db 1371 TAESGKLTATQTGSSITSSNGQTTLTAKDSSIAGNINAANVTNLTGT----- 1417  
QY 1293 IVATTSSAGSGGVSSIPTNPKR-----ISAAVPSGSAATPTMSENKVFEL 1339  
Db 1418 --LTTGDSKINATGTLTINAKDAKLGAAGSDRFTVFNATNAGSGNVTAKTS--SSVNI 1474  
QY 1340 TGDLTLDIPNGNFYQN-----PWLGSDDLVDPLIK 1368  
Db 1475 TGDLTNGLNIISGNRTVRLRGKEIDVKYIQ 1508

RESULT 12

US-08-728-470-9  
; Sequence 9, Application US/08728470  
; Patent No. 5928651  
; GENERAL INFORMATION:  
; APPLICANT: Barenkamp, Stephen J  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; STREET: Bldg. 1  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/728,470  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/302,832  
; FILING DATE: 16-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US93/02166  
; FILING DATE: 16-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9205704.1  
; FILING DATE: 16-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berktresser, Jerry W  
; REGISTRATION NUMBER: 22,651  
; REFERENCE/DOCKET NUMBER: 1038-633  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 415-0813  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1338 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-728-470-9

Query Match 4.1%; Score 366.5; DB 2; Length 1338;  
Best Local Similarity 21.0%; Pred. No. 5.2e-16;  
Matches 257; Conservative 158; Mismatches 465; Indels 345; Gaps 52;

QY 36 SSSFTTETIGRAGAEYIVSGNASFTKFTNIPPTDTTTTPTNS-----NSSSSSG 84  
Db 217 TSLTTLTNT-----TISNLLKSAHVVNITARRKLTVNSSSISIERGSHLILHSEUQGG 268  
QY 85 EFASVSEDSSTTTTPDPKGG-----GAPYNAH-----SGVLSFWTRSGT---EGLSLT 129  
Db 269 QGVQIDKDI TS-----EGGNLTIIYSGWVDVHKHNTILGSGFLNITTKREGDAFEDKSG 321  
QY 130 LSEIKMTGEG-----GAIESQ-----GELLTDL---TSLTIONNLSQLSGAIF 171  
Db 322 RNNLTITTAOGTTTSGNSNGFRFNNVSLNLSGGKLSFTSDREDGRRTKGNISKNKFDG--- 378  
QY 172 GGSTISLSGITTKATFSCNSAEVPAPVKKTEPKAQTASETSGSSSSGSDNSVSPSSSRA 231  
Db 379 ---TLNISCTVDISMKAPKVSFWYRDGRTYWNVVTLNVTSGSKF---NLSIDSTSGST 432  
QY 232 EPAANLQSHFICATATPAAQDTETST-----PSHKFGSGGAIYA--KGDUTIAISQE 283  
Db 433 GPSIRNAELNGI--TFNKATFNIAQGSTANFSIKASIMPFKSNANYALFNEDISVSGGS 490  
QY 284 VLFSINKATKD--GCAIFAEKDVSPENITSLKVOTNGAEKGGAIYAKGDLISQSKQS 340  
Db 491 VNFKLMASSNIQTTPGVIIKSNQFNVSGGSTLNLKAEGSTETAFSI--ENDLNLAI--- 545  
QY 341 LFNSNYSKGGGALYVEGGINFQDLEIRIKYNKAGTFTETKTLPLSLKAQASAGNADAW 400  
Db 546 -----GGNITTRQVEGTSRVNK-GVAAKKNTF-----KGGNITFG 581  
QY 401 ASSPSGSGATTVSDSGSDSGSDSTSETVPVTAAGGLYTDKNLSITNITGIIIEIAN 460  
Db 582 SOKATTEIKGNVTINKNTNA-----TLRGANF--AENKSPLNIAAGNV-INN 624  
QY 461 NKATDVGGGAYVKGTLTCEHSHLQFLKN-----SSDKQG-----GGIYGED--- 502  
Db 625 GNLTTAGSIINIAGNLTVSKGANLQAITNVTNVAAGSDNNGASNTSIARGGAKFKDINN 684  
QY 503 ----NITL-SNLTGKTLFOENTAKEEGG-GLFIKGTDKALTMGTGLDSFCLINNTSEKHGG 556  
Db 685 TSSLNITNTSDTYRTIIGKNISNKSGLDNIIDKSDAEIQIG-----NISQEG- 735  
QY 557 GAFVTKELISQTVTSVDVETIPG-ITPVHGETVITGNKSTGNGGG---VCTKRLLALSNO 611  
Db 736 -----NLTISSDKVNITNQITIKAG---VEGGRSDSSEAEANLTIQTKEKLA3-- 782  
QY 612 SISISGNSAAE---NGGGAHTCPDSFPPTADTAEQPAASAAATSTPKSAPVSTALSTPSS 667  
Db 783 DLNISGFNKAETAKNGSDLTIGNASGGNAD-----AKKVTFDKVKDKSIST----- 829  
QY 668 STVSSLTLLAASSQASPATSNKETQDPNADTDLIDYVVDVTITISKNTAKKGGIYAKAK 727  
Db 830 -----DGHNVTLNSEVKTNGSSNAGNDNSTGLTISAKDVTVNNVNTSHKT- 875  
QY 728 MSRIDQLNISENATEIGGICCKESLELDAL-----VSLSVTE 766  
Db 876 -----INIS-----AAAGNVTTKEGTTINATTSVEVTAQNGTIKGNITTSQNVVTATE 924  
QY 767 NLVKGEGGGLHAK--TVNIS---NLKSGFSFNKANSSTCVATASAPAAAASIAQA 820  
Db 925 NLVTTENAVINATSGTVNISTKTGDIKGGIESTSGMNVNITASGNTLKV5----- 973  
QY 821 AAAAAPSPATPTYSGWVGGAIFYERKVTFSQCSGTQFCQSGNQAIIDNNPSSOSSLNVGGAI 880  
Db 974 -----NITGQDVTYTADAGALTTTAGTISATTNANTTTKTDI 1013  
QY 881 YAKTSLSIGSDAGTSYIFSGNSVSTKQSTGTGQIAGGAIYSPVTVTLNCPATFSNNTASI 940  
Db 1014 NGKVSSSGSV-----TLVATGATLAVGNISG-----NTVTI 1045  
QY 941 ATPKTSSEDSGSGNSIKDITIGGAIAGTATLSG--VSRFSGNTADLGAAG--TLANANT 996  
Db 1046 -----TADSGKLTSTVGSTINGTNSVTTSQSGDIEGTISGNTVNTASTGDLTIGNSAK 1100



Db 1149 -----NVTNLTGTLTTTGD---SKINATSGTLTINAKDAKLDOGAAGD-----RTV 1192

QY 1115 AILTLASSGNITFSNNSLONQGD 1139

Db 1193 VNATNAGSGNVTAKTSSSVNITGD 1217

RESULT 14

US-09-377-155-33

Sequence 33, Application US/09377155

Patent No. 6197312

GENERAL INFORMATION:

APPLICANT: PEAK, Ian Richard Anselm

APPLICANT: JENNINGS, Michael Paul

APPLICANT: MOXON, E. Richard

TITLE OF INVENTION: NOVEL SURFACE ANTIGEN

FILE REFERENCE: 065064/0128

CURRENT APPLICATION NUMBER: US/09/377,155

CURRENT FILING DATE: 1999-08-19

PRIOR APPLICATION NUMBER: PCT/AU98/01031

PRIOR FILING DATE: 1998-12-14

PRIOR APPLICATION NUMBER: GB 9726398.2

PRIOR FILING DATE: 1997-12-12

NUMBER OF SEQ ID NOS: 33

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 33

LENGTH: 2353

TYPE: PRT

ORGANISM: Haemophilus influenzae

US-09-377-155-33

Query Match 4.0%; Score 361.5; DB 4; Length 2353;

Best Local Similarity 19.2%; Pred. No. 2.6e-15;

Matches 344; Conservative 265; Mismatches 684; Indels 499; Gaps 80;

QY 7 TAVFAVL-PSVSGFCFPEPEKLN-----FSRVETSSSTFTTET----- 44

Db 36 TAVLATLLFAVQANATDEDELDPPVPTAPVLSFHSDEKGETEKEVENTSNNGIYFDNK 95

QY 45 -IGEAGAEIVSGN-----ASFTKFTNPTTD-TTTTNSNSSSSGETAS 88

Db 96 GVLKAGAILKAGDNLIKONTDESTNASSFTYSLKDLDTLSVATEKLUSFGANGDKVD 155

QY 89 VSDSDSTTTTPDKGGGFAYNAHSGVLSFMTRSG--TEGSLTLSELKMTGEGGAIFSG 146

Db 156 ITSDANGLAKTNGNVHLNGLDLTPDAVTNTGVLSSSFTPNDEKTI----- 205

QY 147 ELLFTDLTSLTIONNIS---QLSGGAIFGGSTISL-----SGITKATFSCNAAEVPAPV 197

Db 206 -----RAATVKDVLNAGNWKIAGAKTAGGNVESVDLVSAYNNVEFTTGDKNLTDVVLTA 258

QY 198 K---KPTD---PKAQTASSETS---SSSGNDSS---VSSPSSSRABPAAANLOSHFICAT 246

Db 259 KENGKTEVKFTPKTSVIEKDKGLFTGKENDNTNKTSTNTADNTDEGNGLVTAKAVID 318

QY 247 ATPAAQTDFTETSPSHKPSGGGAIYAKGDLTIADSOEVLFSINKATKDGAFAEKD--- 303

Db 319 AVNKAHRVKVTTTANGONGDFATVAGSTNVTTFESGDTTASVTKDT-NGNGIIVKVDKAV 377

QY 304 ---VSFENTISLKVQNGAEKGAAYAKGDLISQSSKOSLFNS-----NYSKQGGG 352

Db 378 GDGLKFDSKKIVADTALTAVTGKV---AEIAKEDDKKLVNAGDLVTLALGNLSWKAKA 434

QY 353 ALYVEG---GINFDLEEIRIKYNKAGTETTKITLPSLKAQASAGNADAWASSSQSGS 409

Db 435 EADTDGALLEGIS---KQDEVKAGETVTFKAGK-----NLKVQDGAN---FTYSLQDALT 483

QY 410 GATTVSDSGDSSSGSDST---SETVPVTAKGGLVTDKNL-----SITNI- 452

Db 484 GLTSLTGLGTNGNDKAVINKDGLTITIPAGNGGTTGTNTISVTKDKGAKAGKAITNVA 543

QY 453 TGI-----IEIANNKATDVG---GGAYVKGTLTC---ENSHRLQFLKNSSDRQGGGIYG 500

Db 544 SGLRAYDDANFDYLNNSATDLNHRVEDAY-KGLNLNEKNANKQPLVTDSTAATVGD-- 600

QY 501 EDNITLNLTKGTLFORNTAKBEGGLFIKGTOKALTMGTGLDSFCLINNTSEKHGGAFV 560

Db 601 -----RKLGVVYSTKNGTKEESNQ--VQADEVL-FTGAGATV---TSKSENGKHTI 647

QY 561 TREISOT-----YTSDEVETIPGVHGETVITGNKSTGGNGGVCYTKRLALSNLQSI 615

Db 648 TVSVAETKADCGLEKGDGTI---KLKVDNQNTDNVLTVCNNGTAVTK----- 691

QY 616 SGNSAAENGGAUTCPDSDPPTAD-----TABQPAASAASTSTPKSAPVSTALSTPSSSTVS 671

Db 692 -----GGFETVKTGATDADRGRKVTYKDATANDADKKVATVKDVATAINS-AATFVK 741

QY 672 SLTLAASSQAAPATSNKETQDPNADTDLLIDYVVDTTISKNTAKKGG-----I 721

Db 742 TENLTTSIDEDNPTDNGKDDALKAGDT-----LTFKAGKNLKVRRDGNITFDLAKNL 794

QY 722 YAKKAKMSRIDQLNISENSATEIGGGTCCKESLELDALVLSVTENLVGKEGGGLHAKTV 781

Db 795 EVKTAKVS--DTLTIGNTPT-----GCTTATPKV 822

QY 782 NISNLKSGFSSNNKANSST-----GVATTASAPAAAASLOA-----AAAAAPSSUATP 832

Db 823 NITSTADGLNFAKETADASGKNVYLKGIATLTLEPSAGAKSSHVDLNVDATKKSNAISI 882

QY 833 TYSGVVGGAIYG-----EKVTFESOCs-GTCOFSGNQALDNNPSQ-----SSL 873

Db 883 EDVLRAGNIOGNGNVVYATYDTVNFDDSTGTTVTVTQKADGKADGVKIGAKTISVI 942

QY 874 NVOGGAIYAKTSLSIGSS-----DAGTSYIFSGNSVS-----TGKSQT--T 912

Db 943 KDINGKLFTCKDLKANDANGATVSEDGCKDFTGLVTAKTVIDAVNKSGRWRTGEGAIET 1002

QY 913 GQIAGGAIYSPVTNLNCPATFSPNTASIAIATPKTSSSEDSGSGNSIKDTIGGAIA----- 965

Db 1003 GATAVAGNAETVTSCTSVNFKNGNATTAT--VSKONGNINVKYDVNVGDLGIGDEKKI 1060

QY 966 ---GTATILSG--VSFRSG-----NTADLGAAGICTLA-NANTPSATSGSONSIT 1008

Db 1061 VADTTTLTVTGKVSVPAGANSVNNKLVNAEGLATALANLNSWTAKADYADGESEET 1120

QY 1009 -----EKITLENG-----SFIFE-----RNQANKRGAIYSPS 1035

Db 1121 DOEVKAGDKVTFRAGRNMLKVQSEKDEFTYSLODTLGLTITLGGTANGRNDTGV--- 1176

QY 1036 VSIKGNITFNQ-----NTSTHDGSAIYFTKDATIESLGSVLFTGN-NVTATQASSATSGQ 1090

Db 1177 INKDGTLITLANGAAAGTASNGNTISVTKDG-----ISAGNKEITNVKSALKTYKD 1228

QY 1091 NWTAN-----YCAAI-----FGDPGTOSQTDAILTLASSGNITFSNNSLQNN 1136

Db 1229 TQNTAEDTQKEHAAVKNANEVEFVGKNGATVSAKT-----NNGKHTVTDVAEAK 1281

QY 1137 QGTPASKFCFSIAGYVKLSLQAAGKTISEFFDCVHTSTKKTGTQNVYETLIDINKEESN 1196

Db 1282 VGDLEK---DTDGKIKLVNDTGNLLVVDATKGASVAKGEENAVTTDTATTAQGTNAN 1338

QY 1197 PYTGTIVFSELHENKSYIPQNAIHLNGTLVLKEKTELHVVSFEQKEGSLIMEPGA/LS 1256

Db 1339 ERGKVVVWVKS-----NGATATETDKKVVATVG---DVAKAINDAATFVK 1379

QY 1257 NONTANGALAINGLITDLDSMGTPOAGEIEFSPPELRIVATTSASGSGSVSSSPTPNPKR 1316

Db 1380 VEN--DDSATIDDSPTDDGANDALKAGDTLT---LK-----AGKNLKVRRDGN 1423

QY 1317 ISAAVPSSGAATPTMSE-----NKVELTGDLTLDIPNGNFYQNPMLGSLDLDVPLIK 1368

Db 1424 ITTALANDLSVKSATVSDKLSLGTNGKNKVNITSDTKCL-----NEAKDSKTGDDANIHLNG 1479

QY 1369 LPTNTSDVOVYDLTSLSGDLFPQKGYMGVGTLLDSNPQTKLOARWTFDYRRWYIPIRINH 1428



```
QY 1036 VSIKGNITFNO-----NTSHDGSALYFTKDATIESLGSVLFTGN-NVTATQASSATSGQ 1090
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1177 INKDGLTITLANGAAAGTDAENGTISVTKDG-----ISAGNKEITNVKSALATYKD 1228
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1091 NNTFAN-----YGAAI-----FGDPCTTOSSTQDAILTLASSGNITFFSNNSLQNN 1136
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1229 TONTADETQDKEFHAAVKANANEVEFKNGATVSAKD-----NNGKHTVTIDVAEAK 1281
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1137 QGDTPAKFCISAGYVKLSLQAAKVTISFFDCVHTSTKKTGSTQNVYETLIDINKEENS 1196
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1282 VDGLEK---DTCGKTKLVNDTNGNLLTVDATKASVAKGEFNAVTTDATTAGGTNAN 1338
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1197 PYTGTVFSELHENKSYIPQNAIHLNGTLVLKEKTELHVVSFEQKEGSKLIMEPGAVLS 1256
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1339 ERGKVVVKGS-----NGATATETDKKKVATVG---DVAKAINDAATFVK 1379
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1257 NQNIANGALAINGLTIDLSSMGTPQAGEIFSPPELRIVATTSSASGSGVSSSIPTNPKR 1316
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1380 VEN--DQSATIDDSPTDDGANDALKAGDTLT---LK-----AGKNLKVKRDGKN 1423
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1317 ISAAVPSGSAATPTWSE-----NKVFLTGDLTLIDPNGNFYONPMLGSDLDVPLIK 1368
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1424 ITPALANDLSVKSATVSDKLSLGTNGKNVNITSDTKGL-----NPAKDSKTGDDANHLNG 1479
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1369 LPTNTSDVQVYDITLSGDLFPQKGYMGWTWLDNSNPOTGKLQARWTFDTYRRWYIIPRDNH 1428
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1480 IASTLTD-----TLLNSGATTNLGGNGITDNEKKRAASVKD-- 1515
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1429 FYANSILGSONSMIVVKQGLINMLNN-----AREDDIAYNNFWVSGVGTPLAQOQTPLSE 1484
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1516 -----VLNAGWNVRGVKPASANNQOVENIDFVATYDTVDF-----VSG-----DKDT---- 1556
: : | | | : : | | | : : | | | : : | | | : : | | |
QY 1485 EFSYYSRGTSAIDAOK---PRODFILGAAFSKIVGKTKAIKKMNYFHKGSE 1533
: : | | | : : | | | : : | | | : : | | | : : | | |
Db 1557 -----TSVTVESKDNKRTEVKIGA-----KTSVIKDHNGKLFYCKE 1593
: : | | | : : | | | : : | | | : : | | | : : | | |
```

Search completed: October 25, 2001, 08:41:01  
Job time: 350 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 25, 2001, 08:39:46 ; Search time 16.37 Seconds  
(without alignments)  
3666.195 Million cell updates/sec

Title: US-09-454-684A-180

Perfect score: 8942

Sequence: 1 MKWLSATVFAVLPVSGF.....IDVGMVTLQMTSCGARMIF 1752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	452	5.1	2249	1 190K_RICRI	P15921 rickettsia
2	422	4.7	2003	1 YBBA_ECOLI	P33666 escherichia
3	397.5	4.4	1656	1 OMPB_RICJA	Q06653 r outer mem
4	395.5	4.4	1140	1 YH96_YEAST	Q04893 saccharomyc
5	394	4.4	964	1 OMPB_CHLPR	O84877 chlamydia t
6	371	4.1	1643	1 OMPB_RICPR	Q53020 r outer mem
7	371	4.1	1654	1 OMPB_RICRI	Q53047 r outer mem
8	359.5	4.0	1034	1 OMPF_CHLPR	P38008 chlamydia t
9	353.5	4.0	1306	1 MSB2_YEAST	P32334 saccharomyc.
10	346.5	3.9	1367	1 AMYH_YEAST	P08640 saccharomyc
11	342.5	3.8	1609	1 FIG2_YEAST	P25653 saccharomyc
12	341.5	3.8	1567	1 ICEN_XANCT	P18127 xanthomonas
13	339.5	3.8	1645	1 OMPB_RICTY	P96989 r outer mem
14	334.5	3.7	1848	1 CBPA_CLOCL	P38058 clostridium
15	325.5	3.6	1150	1 AFMO_PIG	P12021 sus scrofa
16	321.5	3.6	1300	1 120K_RICRI	P14914 rickettsia
17	316.5	3.5	3591	1 FHAB_BORPE	P12255 bordetella
18	315	3.5	1802	1 HKR1_YEAST	P41809 saccharomyc
19	314	3.5	1608	1 HLYA_SERMA	P15320 serratia ma
20	309	3.5	1161	1 YJ9P_YEAST	P41799 saccharomyc
21	307.5	3.4	995	1 YI09_YEAST	P40442 saccharomyc
22	305	3.4	881	1 YJH8_YEAST	P47033 saccharomyc
23	305	3.4	2132	1 PGCA_MOUSE	Q61282 mus musculus
24	304.5	3.4	2124	1 PGCA_RAT	P07897 rattus norv
25	301	3.4	2090	1 N214_HUMAN	P35658 homo sapien
26	299	3.3	1286	1 AIDA_ECOLI	Q03155 escherichia
27	293	3.3	1250	1 YFAL_ECOLI	P45508 escherichia
28	293	3.3	1577	1 HLYA_PROMI	P16466 proteus mir
29	290.5	3.2	2329	1 YS89_CAEEL	Q09624 caenorhabdi
30	288.5	3.2	1569	1 YFJA_ECOLI	P52143 escherichia
31	287.5	3.2	1251	1 Y0U3_CAEEL	Q09550 caenorhabdi
32	284	3.2	959	1 N100_YEAST	Q02629 saccharomyc
33	281.5	3.1	2334	1 WAPA_BACSU	Q07833 bacillus su

## ALIGNMENTS

RESULT 1  
190K\_RICRI  
ID 190K\_RICRI STANDARD; PRT; 2249 AA.  
AC P15921;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE 190 KDA ANTIGEN PRECURSOR (CELL SURFACE ANTIGEN).  
OS Rickettsia rickettsii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=783;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R;  
RX MEDLINE=90354033; PubMed=2117568;  
RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;  
RT "A protective protein antigen of Rickettsia rickettsii has tandemly repeated, near-identical sequences.";  
RL Infect. Immun. 58:2760-2769(1990)  
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.  
CC -!- PTM: GLYCOSYLATED (POSSIBLE).  
CC -!- DISEASE: PROBABLY PLAYS AN IMPORTANT ROLE IN THE PATHOGENESIS OF AND IMMUNITY TO ROCKY MOUNTAIN SPOTTED FEVER.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M31227; AAA26380.1; -.  
CC PIR; A41477; A41477.  
CC Antigen; Repeat; Signal; S-layer; Glycoprotein.  
CC SIGNAL 1 28 POTENTIAL.  
CC CHAIN 29 2249 130 KDA ANTIGEN.  
CC DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.  
CC REPEAT 212 286 A (TYPE I).  
CC REPEAT 287 358 B (TYPE II).  
CC REPEAT 359 430 C (TYPE II).  
CC REPEAT 431 505 D (TYPE I).  
CC REPEAT 506 577 E (TYPE II).  
CC REPEAT 578 652 F (TYPE I).  
CC REPEAT 653 724 G (TYPE II).  
CC REPEAT 725 799 H (TYPE I).  
CC REPEAT 800 874 I (TYPE I).  
CC REPEAT 875 949 J (TYPE I).  
CC REPEAT 950 1021 K (TYPE II).  
CC REPEAT 1022 1093 L (TYPE II).  
CC REPEAT 1094 1165 M (TYPE II).  
CC REPEAT 1166 1180 TYPE I (INCOMPLETE).  
CC SEQUENCE 2249 AA; 224333 MW; A9D6646C089DF087 CRC64;  
SQ

34 276 3.1 1025 1 SLAP\_CAUCR  
35 275.5 3.1 1419 1 ALA1\_CANAL  
36 273 3.1 725 1 AGA1\_YEAST  
37 269 3.0 1075 1 FLO5\_YEAST  
38 269 3.0 1537 1 FLO1\_YEAST  
39 266.5 3.0 1200 1 ICEN\_PSESY  
40 266 3.0 1113 1 N116\_YEAST  
41 265 3.0 1196 1 ICEV\_PSESY  
42 264 3.0 1120 1 STFR\_ECOLI  
43 261.5 2.9 1169 1 YK82\_YEAST  
44 261 2.9 1041 1 EGT2\_YEAST  
45 260.5 2.9 2090 1 HFC1\_MESAU

P35828 caulobacter  
O13368 candida alb  
P32323 saccharomyc  
P38894 saccharomyc  
P32768 saccharomyc  
P06620 pseudomonas  
Q02630 saccharomyc  
O33479 pseudomonas  
P76072 escherichia  
P36170 saccharomyc  
P42835 saccharomyc  
P51611 mesocricetu

Query Match		5.1%; Score 452; DB 1; Length 2249;
Best Local Similarity		20.1%; Pred. No. 2.2e-13;
Matches		445; Conservative 274; Mismatches 822; Indels 678; Gaps 99;
Qy	29	NFSRVETSSSTFTTETGEAGAEEIVSGNASFTKFTNPIPTDTPPTNSNSSSSGETAS 88
Db	214	NTNALATVNVGAGTATLGGAVIK-----ATTTLTN--AASVLTLTN-----AV 257
Qy	89	VSEDSSTTTTDPKGGAGAFYNAHSGTSLTSEIKMTGEGGAIFSGEL 148
Db	258	LTGAIDNTT-----GGD-----NVGVNL-----NGALS-----QVTGDIGN----- 289
Qy	149	LFTDLTSLTIQNNLSOLS--GGAIFGSGTISLGIKATPSCNSAEVAPVKKPTEPKA 205
Db	290	-----TNSLATISVGAGTATLGGAVIKAT-TTKLTDAASAVKFTNPV-VVTGAID 337
Qy	206	QTASETSGSSSSGNSVSSPSSRABPAAN-----LQSHFICATATPAAQTDTERSTP 260
Db	338	MTGNANNGIVFTTGNSTVTVGNVTNALATVNVGAGLLQVGGVVKANTINLTDNASAVT 397
Qy	261	SHKP-----GSGGAIYAKGDLTIA--DSQEVLPFSINKATKD---GGAIFAEX 302
Db	398	PTNPVVTGADNTGNANNGIVFTTGNSTVTDIGNTNALATVNVGAGTATLGGAVIKAT 457
Qy	303	DVSPENITSLKVQTNGAEEKGAIIYAKGDLISQSSKOSLFNSYSKQGGALYVEGGI-N 361
Db	458	TTKLTNAASVLTLTNANAVLTGAI-----DNTTGDVNVGLNLN-----GALSQVTVGNIGN 508
Qy	362	QDLLEEIRIKYKAGTETTKITLPSLKAQASAGN----ADAWASSPQSGSGATTVSDS 417
Db	509	TNSLATISV---GAGT-----ATLGGAVIKATTTKLTDAASAVKFTNPVVTGTA--IDNT 558
Qy	418	GDSSGSDSDTSEVPVTKAGGGLYTDKNLSITNITGIEIANKKATDVGGAVVKGLT 477
Db	559	GNANGI-----VFTTGNSTVTC---DIGNTSLATISVGAGTATLGGAVIKATTT 606
Qy	478	CENSHRLQFLKNSDKQGGI---YGEDNITLNLATGK-----TLFOENT 519
Db	607	KLTNAAASVLTLTNANAVLTGAI-DNTTGDVNVGLNLNALSQVTDIGNTNSLATISVGA 666
Qy	520	AKESGGGLFKGTOKALT-----MTG-LQSFCLINNTSKHGGGAPVTKEIS 565
Db	667	GTATLGGAVIKATTTKITNAVSARKFTNPVVTGAI-DNTTGDVNVGLNLNALSQVTDIGNTNS 726
Qy	566	QTYT-----SDVETIPG-----ITPVHGETVITG--NKSTGSGGGV 600
Db	727	NTNALATVNVGAGTATLGGAVIKATTTKLTNAASVLTLTNANAVLTGAI-DNTTGDVNVGL 786
Qy	601	CKRLALSNLQSIISGNSAA--ENGGAHTCPDPTADTAEQPAAASATSPKSPAPV 658
Db	787	LNALGALSQVTDIGNTNSLATISVGAGTATLGGAVIKATTTKLTNAASVLTLTNANAVL 846
Qy	659	STAL-STPSSSTVSSLLLAASSQAASPATSKETQDPNADTDLIDYVDTTISKNTAKK 717
Db	847	TGAVDNTTGDVNVGLNLGALSQVTDIGNTNS-----LATISVGAGTATL 893
Qy	718	GGG-IYAKKAKMSRIDOLNISENATELGGGICCKESLELDALVSLVTENLVKKEGGL 776
Db	894	GGAVIKATTTKLTNAASVLTLTNANAVLTGAI-----DNTTGDVNVGL 936
Qy	777	HAKTVNISNLKSGFSFNKANSST-----GVATTASAPAAAAAS--LQAAAAAPSSP 829
Db	937	-----LNLGALSQVTDIGNTNSLATISVGAGTATLGGAVIKATTTKLTDAASAVKFTNP 992
Qy	830	ATPIYGVVGAIGYKVTFFQCSGTCOFSCNOAIDNNPSSO-----LNVQGG 878
Db	993	V-----VVTGAIDN---TGNANNGIVFTTGNSTVTVGNVTNALATVNVGAGLLQVGG 1043
Qy	879	AIYAKTSLSIGSDAGT-----SYIFGNSVSTGKSQTTG----- 913
Db	1044	VKANTINLTDNASAVFTNPVVTGAI-DNTGNANNGIVFTTGNSTVTVGNVTNALATV 1103
Qy	914	-----QIAGGAIYSPTVTL--NCPA-TFSN-----NTASTATPKTSSD 949

Db	1104	NVGAGLLQVQGVVVKANTINLTDNASAVFTNPVVTGAI-DNTGNANNGIVFTTGNSTVT 1163
Qy	950	GSSNSIKDTIGGAGIAGTATTLTSGVSRSFGNTADLGA----- 986
Db	1164	GDIGNT--NALATVNVGAGITLQAGGSLAANNIDFGARSTLEFNGLDGGKAIPYFVK 1221
Qy	987	-----AICTLANANTPSATSSQNSITEKITLENGSIFERNQ 1024
Db	1222	AIANGNAILNVNTKLTITASHLTIGTVAEINIGAGNLFTIDASVGDVITLNAQINFPAR 1281
Qy	1025	-----ANKRGA-----IYSPSV-----SIKNNITFN-----QNTSTHDG 1054
Db	1282	DSVLVLSNLGTGVGNILLAAADLVAPGADEGTVVFGVGNVGLNVSNAVAGTARNIDGGG 1341
Qy	1055	-----SAIYTKDATTESLGSVLTGN-----NVATQASSATSGQNTTAN 1096
Db	1342	NKFNLLIYNAVTTTDDVNLGEGIQNLINKNADEFTSSTAFNAGAIQINDATYTTDANNGN 1401
Qy	1097	YGAIFGDPGTTQSSQTDAILTLASSGN--ITFSNNSLONNOGD-----TPASKFC 1146
Db	1402	LNI-----PAGNIQFAHADAQLVLQNSGSDRTIILGANIDPDNDEGIVILNSVTAGK 1457
Qy	1147	SIAGYVLSLQAARKTISF---FDCVHTSTKTKTGSTQNVYETLIDINKEENSNPYGTIV 1203
Db	1458	TIAG-GKTEFGGAHLQITILPKGAGDCSTAGT--TFNTTNI--VLDITGOLELGATTANVV 1512
Qy	1204	FSSE-----LHENKSY---IPONAILHNGTLVLKKEKTELH- 1235
Db	1513	LFNDAVOLTOFTGNIGGFLDFNAKMGVTLNNNVNAGAVQNTGTTNNGTTLVLGASLN 1572
Qy	1236	---VVSFEQKEGSKLIMEPGAVLSNQ-----NIA 1261
Db	1573	VNGIAMLKVAGAGNTIAKGGVKIGEIOGTGTNTLTLPAHFNLTGSINKTGOALKLFM 1632
Qy	1262	NGALAINGL-----TIDLSSMG-TPQAGEIFSPPELRIIVATTSSASGGSVSS----- 1308
Db	1633	NGG-SVSGVVGTAANSVGDITTAGATSFASVNAKGTATLGGTTSFANTFTNTGAVTLAK 1691
Qy	1309	SITNPKRISAAVPSGSA-----TTPTMSENKVFLTG-----DL 1343
Db	1692	GSITSFAKNVTATSFVANSATINFNSLAFNSITGGGTTTLTGANQVYTGTSFTDTL 1751
Qy	1344	TL-----IDPNG-----NFQNPMLGSDLDVPLKL-- 1369
Db	1752	TLNTTFDGAASGGNLIKSGSTLDLSCVSTLALVVTATNFDMN-NISPDKIYFVISET 1810
Qy	1370	-----PTNTSDVQV-----YDLT-----LSGDLFPQKGYMGWTWL 1399
Db	1811	AGGLKPTSKENVKITINNDNRFDVTFDASTLTTLFAEDIAADVIDGDFAP-----GGPL 1864
Qy	1400	DSNPQTGKLO-----ARWTFDYRRVVIYPR---DNHFYANSILGSONSMI 1442
Db	1865	ANIPNAANIKSLELMEDAPNGSDARQAFNNFGLMTPLOEADATTHLIQDVVKPSSDTIAA 1924
Qy	1443	VVKGLINNMNLN-----NAREDDIAYNN-----FWVSGVTFLLAQOQTP 1481
Db	1925	VNNQVVASNISSNITALNARMKQVSGNKGVPSSGDEMDAKFGAWIS---PFVGNATQK 1981
Qy	1482	LSEBSFY---SRGTSVAIDAKPRODFILGAFAFKIVGKTKRAIKKMHNYFHKGSEYSQA 1538
Db	1982	MCNISISGYKSDTTGTTIGDFGVSDDLALGLAYTR---ADTDIKLKN--KTGOKNKVES 2036
Qy	1539	SVYGGKFLYFLNKHQHWALPF---LIQGVVSYGHIKHD-----TTLYPSIHERNKG 1588
Db	2037	NIYS---LYGLYN-----VPIENLFVEAIASYNKIRSKRRVIATTLTVGTGTANG 2087
Qy	1589	DWEDLGLWADLRISMDLKEPSKSDSKRITVYGELEYSIROKQFTEIDYDPRHFDCCAYR 1648
Db	2088	KYKSESYTQGLMAGYTYMMP---ENINLTPLAGLRYSTIKDKYKE-----TGTYIQ 2136
Qy	1649	NLSLP-----VGAIVEGAIMCNCILMYNKLALAYMPSIYRNPNVCKYRVLSSNEA 1698



Db 2137 NUTVKGKNTNFDGLGAKVSSNI-NVNEIVLTPELYAMVDYAFKN-----KVSADARLQ 2191  
QY 1699 GVQICVPTRTSARAEYSTQLYLG-----PFWTLXGNYTIDVGMVTLTQMTSCGARMIF 1752  
Db 2192 G-MTAPLPNSFKQSTSDVGVGTAKHKMMEYRINYDTNIGSKYFAQGGVKYRVNF 2249

## RESULT 2

YDBA\_ECOLI STANDARD; PRT: 2003 AA.  
AC P33666; P76087; P76088; P76856; P76857; P76859;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DE 01-OCT-2000 (Rel. 40, Last annotation update)  
DT HYPOTHETICAL 205.9 KDA PROTEIN IN PAAY-ACPD INTERGENIC REGION.  
GN YDBA OR B1401/B1405.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_Taxid=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=97251357; PubMed=9097039;  
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,  
RA Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,  
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
corresponding to the 28.0-40.1 min region on the linkage map.";  
RL DNA Res 3:363-377(1996).  
RN [3]  
RP SEQUENCE OF 464-2003 FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=92190338; PubMed=1665988;  
RA Moszer I., Glaser P., Danchin A.;  
RT "Multiple IS insertion sequences near the replication terminus in  
Escherichia coli K-12.";  
RL Biochimie 73:1361-1374(1991).  
CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).  
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. THE GENE CODING FOR  
THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT  
BETWEEN AMINO ACIDS 839 AND 840.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000237; AAC74483.1; ALT\_SEQ.  
DR EMBL; AE000237; AAC74487.1; ALT\_SEQ.  
DR EMBL; D90778; BAA15009.1; ALT\_SEQ.  
DR EMBL; D90778; BAA1880.1; ALT\_SEQ.  
DR EMBL; D90779; BAA18881.1; ALT\_SEQ.  
DR EMBL; X62680; -; NOT\_ANNOTATED\_CDS.  
DR EcoGene; EG11307; ydbA.  
KW Hypothetical protein.  
FT CONFLICT 489 489 I -> V (IN REF. 2).

FT CONFLICT 495 495 I -> V (IN REF. 2).  
SQ SEQUENCE 2003 AA; 205949 MW; B83A12C8B53220EE CRC64;  
  
Query Match 4.7%; Score 422; DB 1; Length 2003;  
Best Local Similarity 19.7%; Pred. No. 4.6e-12;  
Matches 394; Conservative 245; Mismatches 757; Indels 600; Gaps 86;  
  
QY 28 LNFSRVETSSSTFTTETIGEAGA-----EYIVSGNASFTKFTNIPTD-----TTPP 74  
Db 184 INLWQIDEANNVVALEGVADGATKWOYNHNGELVITGDNAVNNNGKTVTDGKDSITGTE 243  
QY 75 TNSNSS-----SSGETASVSESDSTTTTDPKGGAGAFYNAHSEVL- 116  
Db 244 INGNNGKVIQDGLDVGSGGHGIDITGDSATV--DNKGTMTVTDPSMGIDQDKAIVN 301  
QY 117 ----SPMTRSGT-----EGSLTILS--- 131  
Db 302 NEGESTITNGTGTQINGDDATANNNGKTVYDGDKDSGTGTEINGNNGKVITQDGLDVSOGG 361  
QY 132 -EIKMTGEGGAIFSQGELLFTDLTSLTIQ-----NNLSQSGGAIFGGSTISLSGIT 182  
Db 362 HGIDITGDSATVDNKGKMTMTVDPESIGIQVDGDOAVNNEGE---SAITNGGTGTQINGD 418  
QY 183 KATFSCNSAEVPAPVKKPTPEKQAQTASETSGSSSSGNDVSSPSSRAEPAAANLQSHF 242  
Db 419 DATANNNG-----KTTVDGKDSGTCTEAGNNGKVIQDGLDVSOGG-----3HG 461  
QY 243 ICATATPAAQTDTETSTPSHKPGSGGAIYAKGDLTTADSQEVLFSINKATKGGGAIFAEK 302  
Db 462 IDITGDSATVDNKGKMTMTVDPESIG--IQIDGQATVNNREGESTITNGGT---GTOINGN 516  
QY 303 DVSFENITSLKVQTNGAEEKG-----GAIYAKGDLSSQSSKOSLFNSYKSGGGALY 355  
Db 517 DATANN--SGKTVYDGDKDSGTGKIAGNIGVINDGSLTGTGGAHGVENIG----DNGTVN 570  
QY 356 VEGGINFQDLEERIKYKAGTFETKKITLPSLKAQASAGNADAWASSPQSGGATTVS 415  
Db 571 NKGDIVVSDTSGIVLINGEG-----ATVSNITGDVNVSNATGFSITNS 615  
QY 416 DS-----GDSSSGSD--SDTSETVPVTAAGGGLYTKNLSITNITGIIETANNKATD 465  
Db 616 GKVSLAGSMQVGDFTGVDLNGNNSVTLLAAK-----DLKVVGOKATGINSVGDANTVN 669  
QY 466 VGGGAYVKGTLTCENSHRLQFLKNSSDKQGGIYGED-NITL-----SNLTGKTLF 515  
Db 670 IYGNVLVDKDKTADNAAYEFF----DPSVGINVYGDNDNVTLDGKLTVYSDSEVTSRQSN 725  
QY 516 QENTAKEEGGGLFIKGTDKALMTGDLDSFCLINNTSEKHGGGAFVTKETISQTVSDVETI 575  
Db 726 LFDGSAEKTSGLVVIGDGNVTNNNG--GLELIGEKNALADGSOVTSRLTGYSVTS-VVVV 782  
QY 576 PGITPVH--GEVITCNKSTG-----GNGGGVC-----TRIAL 607  
Db 783 SGESSVYLNGDITISGEFFLGFAGVIRVQDKALLEIGSGATLTMQDIDSEFHHGTVEI 842  
QY 608 SNLQISISG-NSAANGGGAHTCPDSTFPTADTAEP-----AAASAATSTPKAP 657  
Db 843 QNLGFAFVIGENTGINSITSLQNG---KDPAPSPVILLATNGGSATNAGTITKGVTE 899  
QY 658 VSTALSTPSSST-----VSSLTLLAASSQASPATSNKETQDPNADTDLIDY---- 704  
Db 900 QHSVFENKYSTGTSNPSIFNNDVSSITGLVAQSNSTII-----NTDSGIIDLXGKS 950  
QY 705 VYDTTISKVTAKGGGIYAKKAKMSRID---QLNISENSATEIGGCCICKESLELDALVS 761  
Db 951 VGLMLAIADSTAENQKGTITLDSMWVDANDTTAMRDIASNAIDFGTGV----GVGTDSYSG 1006  
QY 762 LSVTENLVGKEGGLHAKTVNTSNLKS---PSFSNKNKANSSTGVTATASAPAAAAASL 818  
Db 1007 AGKNATAINQLGG-----VITYINAGMAAYGASVTINQGTINLEKNGNYDDSLAPNT 1061  
QY 819 QAAAAAAPSPATPTYSGVVG-----GAI--YGEKVTFSQCSGTCQFSCNQ 862

```
Db 1062 LVGMAYEHGTAINDQGVININVTGQAFYNDGTGTIVNVGTCTF-----GVQC-SCNE 1116
Qy 863 AIDNNPSSSLNVGGAIYAKT--SLSGSDAGTSYIFGNSYSTGK----- 908
Db 1117 -YNNDDFTSLIYTGDDFTIRSETVTLNKAAYTDKL-AGNVNSGTLSDQDITVSSGL 1174
Qy 909 -SQTTG-----QIAGGAT-----YSPVTLCNPATFS--N 935
Db 1175 LENTSGGLIINLVKDKAVIKNAGVMTNNVDVSGGILNAGEMTAQITMAGADSSLVN 1234
Qy 936 NTASIAIPTKTSSEDSGSGNSIKDTIGCAIAGTAITLSCVSRFSQNTADL---GAAIGTLA 992
Db 1235 NTGTI--NKIVONAGVFNS-----GSVTGRMSAGGV--FNNQTDGAIMRGAALTGTA 1284
Qy 993 NAN---TPSATSGSONSITEKITLEN-----CSFIFERQO-----ANKRGAIIYSPSVSI 1038
Db 1285 VANNEGTVNLGSSSEGNNTGMLVNNNSAFNRRGEIILDNDKNVHINQSGTLY----- 1338
Qy 1039 KGNITTEONTSTHDSAIY-----FTKDATIE-SLGSVLFTGN----- 1076
Db 1339 --NTGHMISNSSHNGAVNMGGNGRFINDTIDVSAKSLVVSANNAGDQNAFTWNOGD 1396
Qy 1077 -----NVTATQASSATSGQNTTANYGAAIFGDPGTTQ 1109
Db 1397 VINFDHDSASAVKVTHSNFIQAQNDGIMNISCTGAVAMEGDKNAQLVNNGTINLGTAGTTD 1456
Qy 1110 SSQTDAILTLASSGNITFSNNSLQNOGDTPASKFCFSIAGYVKSLSQAAGKGISFPDC 1169
Db 1457 TCMIGMQIDANATADAVIENNTINIFANDSFVSLGTGVGHV-----VNNGTVVITADG 1510
Qy 1170 VHTS--TKKTGTQNYETLNDIKENSNPYGTGTVFSELSHENKSYIPONAILHNGTLVL 1228
Db 1511 VTGSLIKQGSINW-----EGMNGNGN---SEVHYGYDTLPD-----VP 1549
Qy 1229 KEKTELHVVSFEQEGSKLIMEPAVL--SNQIANGALAINGLTILSSMGT----- 1279
Db 1550 KPNT--VSTGSGDEAGGSMNLLNGYVVGTVNNGSAGKLKVNNASMNGVEINTGTAGTAD 1608
Qy 1280 -----PQAGEIFSPELRIVATVSSASGSGVY-----SSIPTN 1313
Db 1609 TTVSFDNVVGSNLTDAITAITSVVTAKGSTDASGNDVMTSMKNAYTDVATDASVNDI 1668
Qy 1314 PKRISAAVP-----SGSAATT-----PTMSNKKVFLTGD----- 1342
Db 1669 AKALDAGYTNNELETSLVNGVTAELNSALKQVSGSQATTVFEARVLSNRFSLMLADAAPK 1728
Qy 1343 -----LTLIDPNGNFYQNPMLGSDLDVPLIKLPTNTSDVQYIDL-----TLSGDLF 1388
Db 1729 VNGLIAFNVAAGDPRAELGNN---TEYDMLARKTIDLSESQTSMSLEYGIARLDGD-- 1782
Qy 1389 PQGYMGCTWTLDSNPOTG-----KLQARWTFDYYRRWVYIPRDNHFVANSILGSONSMI 1442
Db 1783 -----GAQKAGDNGVTGYSQFTGLKHOMSPFNGMNN-----NNALRYDVHNL 1825
Qy 1443 VVKOGLINMLNNAARFDDI--AYNNFWVSGVGTFLAOG-----TPLSE-----EFSYYS 1490
Db 1826 DSSRSIAGNTNKTADTDVKQOYLEFRREGAKTFEPSGLKVTYPYAGVKLRHTELEGGVQE 1885
Qy 1491 RGTSTVAIDAKPRODFILGAAPFSKIVGKTAKIKKMHNYPHKGEYSYQASV-----YCGKF 1545
Db 1886 RNAG-----DFNLN-----MNSGETAVDSIVGLKLDYACK- 1916
Qy 1546 LYFLLNKOHGHALPLOGVVSYGHIKHDTTLYPSIHER--NKGDWEDLGLWDLRLISM 1603
Db 1917 -----DGMASATLEGGPNLSYAKSQRTASLAGAGSQHFNVDDGQKGGGINSL-TSV 1967
Qy 1604 DLKEPSKDSKRITVY 1619
Db 1968 GVKYSSKESLNLDAY 1983
RESULT 3
```

```
OMP_RICJA
ID OMP_RICJA STANDARD; PRT; 1656 AA.
AC 00653;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OUTER MEMBRANE PROTEIN B PRECURSOR (168 KDA SURFACE-LAYER PROTEIN)
DE (SURFACE PROTEIN ANTIGEN) (CELL SURFACE ANTIGEN) (SCA5) (ROMP B)
DE (CONTAINS: 120 KDA SURFACE-EXPOSED PROTEIN (SURFACE PROTEIN ANTIGEN)
DE (120 KDA OUTER MEMBRANE PROTEIN OMPB); 32 KDA BETA PEPTIDE].
GN OMPB.
OS Rickettsia japonica.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=35790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YH;
RA Uchiyama T.;
RT "Sequencing of the gene encoding the protein rOmp B of Rickettsia
RT japonica.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC SIMILARITY).
CC -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC LAYER WITH HEXAGONAL SYMMETRY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB003681; BAA20138.1; -
KW Antigen; S-layer; Cell wall.
FT CHAIN 1 1338 120 KDA SURFACE-EXPOSED PROTEIN.
FT DOMAIN 1339 1656 32 KDA BETA PEPTIDE.
FT CHAIN 528 533 POLY-GLY.
SQ SEQUENCE 1656 AA; 168097 MW; 3132A69C9DD5999F CRC64;

Query Match 4.4%; Score 397.5; DB 1; Length 1656;
Best Local Similarity 20.0%; Pred. No. 4.9e-11;
Matches 376; Conservative 252; Mismatches 716; Indels 533; Gaps 91;

Qy 2 KWLSA---TAVFAAVLPVSVCFCPEPEKELNFSRVETSSSTT-----PTEIIGAGAEYI 53
Db 10 KLISAGLVTAATASTIVASFAGSAGMAATQON--RTNCGVATTVGVGVGDQTV--ALANVA 65
Qy 54 VSGNASFTKFTNIPTDTPPTNSNS-----SSSGETASVSDSDSTTTTDPKGGGAFY 109
Db 66 VAPNAVITANAN-NGINLNTFAGSENGFLSNANNAVTVSED-----TTL-----GFIN 114
Qy 110 NAHSGVLFSMFTRSCTEGSLTL---SEIKMTGEGGAIFSGELLFTDLFSL---TIQNMLS 163
Db 115 NAANNANRF-----NLTLDAGKTLTITGG-----ITNVQSAATHNAQNIVA 156
Qy 164 QLSGGAIFGGSTISLSGITKATFSCNSAEVPAPVKKPTPEK-----AQTASETSSSSSS 218
Db 157 KFNCGAATANN--DLISGLGTTDFGAASATLVFDLANPTQKAPLILADNALIVNGANGTL 214
Qy 219 G--NDSVSSPSSRAEPAAANL---QSHFICATATPAAQTDTTSTPSHKPCSGGAIVAK 273
Db 215 NVTNGFIQVSDKSFATYKAINIGDQGFMTNATNALNLQ-----AGGTTINFN 266
Qy 274 GDLTIADSQEVLFSINKATKD-----GGAIFAEKOVSFENITSLKVO----- 315
```

Db 267 G---TDGTRGLVLLSKNGAATDFNVGTSGGNNL---KGIIELNTVAINGOLIANAGPANAV 321  
QY 316 ---TNGAEKGG-----GAIYAKGDLISQSSKSLFNSNSYKQGGGALYVE 357  
Db 322 ICTNNGAGRAAGFVVSVDNGKAATIDGQVYAK-DWVIQSA-----NAN----- 363  
QY 358 GGINFQDLLEIRIKYNKA-----GTFFE-----TKKITLPSLKAQASAGNA 397  
Db 364 GQVNRFRHVDVGIDGTTAFKTAASIVAITQNSNFGTDFGNLAAQVTPDVTMTLTGNETG 423  
QY 398 DAWASSPOSGGATVSDSGD-SSGSDSDTSEVPVTA-----KGGGLYTDKNL 447  
Db 424 DA---NNPENTAGVITFAANGTFLASADANAVANNITAIASGVGVVQLSGTHT-AEL 479  
QY 448 SITNITGIETIA-----NNKA---TDVGGGAYVKGTLCENSHRLOFLKNSDKQGGGIYG 500  
Db 480 RLGNAGSVFKLADGTIVNGKVQTVLGVGLAAGAITLDSATI---TGDIGNGGGAA 535  
QY 501 EDNITLSNLTGKTL-----FOENTAKEGGGLFTKGTDKALMTGLDSF 544  
Db 536 LOSITLANDATKTLTGLGANIISANGGTINFOAN-----GGTIKLTSTQNNIIVD----- 585  
QY 545 CLINTNSEKHGGGAFVTKETISQYTSVDETIPGLTPVHCETVIT-----GNKSTGGNGG 598  
Db 586 CDLAITADTG-----VVDASSLTNAQTLTISGTIGIIGANNNTLIGQFNIGSKTTLNGG 640  
QY 599 GYCTKRLALSNLQISISGNSAAENGGAHTCPDSFPTADTAEQPAASAAATSTPKSAPV 658  
Db 641 NVAINELVIGNNSVQFAHT-----NAGG-----KNIVS 751  
QY 659 SPALSTPSSSTVSSLTLLAA-----SSQASP-ATSNKETQDPNADTDLIDYVVD---TTI 710  
Db 674 GKIIENP---VYNNNTTLAAGTNLGSAAENPLAEINFGSKGARADTVLNVGEGVNLVATNI 730  
QY 711 SKNTAKKGGIYAKKAKSRIDQLAISENSATEIGGICCKESLELDALVLSVTENLVG 770  
Db 731 ITTIDNVGSFVF-----NAGG-----KNIVS 751  
QY 771 KEGGGLHARTVNISMLKSG--PFSNNKANSSSTGVATTASAPAAAAAASLQAAAAAPSS 828  
Db 752 GTVGGQGNKFTVALDNGTTVYKFLGNATFNGNTI-----AANSTLQ----- 794  
QY 829 PATPYSGVVGGAIVGEKVTFCQSCQFQFQSNQAIDNNPSSSLNVQGGAIYAKTSLSI 888  
Db 795 -----ISGNYTADFETASADGTGIVEF-----VNTGPINVLNKQAPVYNALKQITV 840  
QY 889 GS-----SDAGTSYIFGNSVSTCKSQSTQGTQIAGAIYSPVTILNCPATFSNNTASIA 942  
Db 841 SGPGNVVYNEIGNAGNYHGAWTDITAFENSS--LGAVLFLPS---GIPNDAGNTIPLTI 895  
QY 943 PKTSEDGSSGNSIKDTI-----GCAIAGTATITLGSVSRFSGNTADLGAAGITLAN 993  
Db 896 KSTVGNETAEGFSVPVSVIVSGVDSVIADQVIGDQNNIVGLG-----LGSDNGIIVN 947  
QY 994 ANTPSATGSSQNSITEKITLENGSIFERNQANKGAIYSPSVSTKGNITNQNTSTHD 1053  
Db 948 AFTLXAGICTINNQGTVTLGG-----VPNTGTVYGLGTGICASK--EKQVTFETD 998  
QY 1054 GSAIYFTRDATIESLGSVLFT---GNNVTATQASA-----TSGQNTNTAN 1096  
Db 999 -----YNNLGNIIATNTINDGVTVTGTGIAAGGIAGTDFDKITLGSVNGNAN 1047  
QY 1097 --YGAIFDPGCTQSSOTDAILTLASSGNITFSNNLSQNNQGDTPASKFCGIAGYVKL 1154  
Db 1048 VRFADGIF-----SNSTSMIVTTKANNGTVYLGNAFVGNIGDS----- 1086  
QY 1155 SLQAAKGTISFFDCVHTSKTKGTSTQNYVETLDINKENSNPYTGTVFVSSSELHENKSY 1214  
Db 1087 -----DTPVASVRFETGSNNGA-----GLKGNIYSOVIDFGT---YNLGI 1122  
QY 1215 IPONAILHNGTILVLEKETEL---HVVSF---EOKEGSKLINEPCAVLSNQNIANGALAING 1269  
Db 1123 VNSNVILGGSTTAINGKIDLLNTLTTFAGGTSTWGNNTSIETTLTLANGNIGHVIA-EG 1181

QY 1270 LTIDLSMGT-----PQAGEIFSPPELRIVATTSSASGGSGVSSSI-----PTNPXR- 1316  
Db 1182 AQVNATTTCTTTINVDONANANFSGTQ-----TYTTLIOGARFNGTLCGPNFTVTSNRF 1236  
QY 1317 ----TSAAVPSGAATPTPMSENKVELTGDLTLDIDNGNFYQNPMGLSDLDVPLIKLPIN 1372  
Db 1237 VNYGLTIRAAANQDYVITRTNNAEN--IVTNDIT-----NSPFGGAPGVGN-----VTTFVN 1285  
QY 1373 TSDVQVY-DLTLSGDLFPQGYMGWTWLTDSNPQTGLQARWTFDYYRRVYIIPRDNHFA 1431  
Db 1286 ATNTAAYNNLLAKNASADSANFVGTIVTDTSAITNAQILDVAKD-----IOAOLGNRLGA 1340  
QY 1432 NSILGQNSMIVVYKOGLINNML--NNAREDDIAYNNFWVSGVGTFLAQOCTPLSEEFY 1489  
Db 1341 LRYLGTPEMVGSEAGAIIPAAVAAGDEAVDNVAY-GIMAKPEYTDHAKSKKGLAGYAK 1398  
QY 1490 SGTSTVAIDAKPRQDFILGAAFSKIVGKTKAIKKMHNYFHKHSE-----YSQASVYGGKFL 1546  
Db 1399 TTGIVIGLDTLANNLMIGAA-----IGITKTDIK-HQDYKKGDKTDVNGFSFSLYGAQ-- 1451  
QY 1547 YFLLNKQHGWLPELQO--VVSYGHIKHDTTTLVPSIH-----BRNKGDWEDLGLWADL 1599  
Db 1452 QIVEN-----FRAQSAIFSLNOVKNSQRYFPDANGNMSKQIAGNYDNMTFEGNL 1503  
QY 1600 RISMDLKEPSKSDSKRITVYGELEY-----SSIROKQFTEIDYDPRHFDJAY 1647  
Db 1504 TVGYDY--NANQGVLTPTMAGLSYLKSSDENYKGTGVVANKQVNS-----KFSD--- 1551  
QY 1648 RNLSLPVGCAGEGATMN 1664  
Db 1552 -RTDLIVGAKVAGGTMN 1567

RESULT 4  
Y96\_YEAST  
ID Y96\_YEAST STANDARD; PRT: 1140 AA.  
AC Q04893;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL I13.1 KDA PROTEIN IN PRES-FET4 INTERGENIC REGION.  
GN YMR317W OR YMR9924.09.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Churcher C.M., Louis E.J., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DDAJ databases.  
CC -!- DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; 254141; CAA90835.1; -.  
DR SGD; S0004936; YMR317W.  
KW Hypothetical protein; Repeat.  
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match 4.4%; Score 395.5; DB 1; Length 1140;

Best Local Similarity 20.6%; Pred. No. 3.7e-11;

Matches 279; Conservative 214; Mismatches 515; Indels 349; Gaps 55;

QY 36 SSSTTFTTIGEGABEYIVSGNASFTKFTNIPTDTTTTPTNSNSSSSSGETASVSED--- 92

```
Db 6 SKSTATTATT-----SHSSTT---TTSSTTSTTTTSTTSTSTKVTTSPEIIV 52
QY 93 SDSYTTTPDPKGGAFYNAHSGVLSFMTSGTEGLSLILSEIKMTGEGGAIFSQGELLFTD 152
Db 53 SSSSTLV-----SSVPEFTSSSSSLSDTIASILSSSESILVSFSSLYTSSD 99
QY 153 LTSLTIONNLSQLSG-----AIFGGSTISLSCITKATFSCNSAEVAPVKKPTEPKAQTA 208
Db 100 ISSYVNDVESTGSPNSYSAALSTNAQLSSSTETDTSISSAI-----QTS 147
QY 209 SETSGSSSSGNDVSSPSSRAEPAANLQSHFICATATPAAQTDTETPSPHKPGSGG 268
Db 148 SPQTSSSNGGSGSEPLGKSSVLETTASSSDT---TAVTSFTTTLTDVSSSPKISSSGS 204
QY 269 AIYAKGDLTIADSOEVLFSINKATKDGAIF-----AEKDVS-----FENITSUKV 314
Db 205 AVTSVG---TTSDAKSEVFS---SSTSDVSSLLSSTSSPASSTISETLPFSSTILSITSPV 260
QY 315 QTNGAEKGGAIYAKGDLISIQSS-----KQSLFNSNYSKQGGGALYVEGGINFQDLE 366
Db 261 SSEAPSATSSVSSEASSTSSVSSEAPLATSSVVSSEAPSTSSVV-----SS 310
QY 367 EIRKYNKAGTFETKIIPLSKAQASAGNADAWASSPQSGGATTVSDSGSSSGSDS 426
Db 311 EAPSTSSVSSEISSTSSVSSEAPLATSSVVSSEAPSS---TSSVSSEISSTTSS 366
QY 427 DTSETVPYTAGGGLYTDKNLINITIGIIEIANNKATDVGGAIVKGTLCENSHRLQF 486
Db 367 SVSSEAPLAT---SSVVSSEAPSTSSVSSEAPSTSSVSSEA---PSSTSSVSSEISS 422
QY 487 LKNS-----SDKQGGGIYGDNTLSNLGKTLFOENTAKEEGGLFIKCTDKALMTWG 540
Db 423 TKSSVMSSEVSATSSVSSSEAPSAISLASSRLFS-----KMTSVTSLIVA 470
QY 541 LDSFCLINN---TSEKHGGGAFVTKETISQYTSQVETIPGTPVHGTEVITGNKSTGGG 597
Db 471 TEASSVTSRLPSPSETLASNIESSLTGYNSTVST---TTSAASSTL---GSKVSSSN- 524
QY 598 GCVCTKRLALNLOSIS-----ISGNSAENGCGGAHTCPDFPADTAEQPAASAA 649
Db 525 -----SRMATSSTSDLSKSVIFGNSST-----VTTSPSASISLTSPLFSVMSDI 574
QY 650 TSTPKSAPVSTAL---STPS---SSTVSSLTL-----AASSQASPATSNKETQDPN- 695
Db 575 TSSSEASS-ISSNLASSAPSDNNSTIASASLIVTKNSVVSSIVSSITSETNESMLA 633
QY 696 -ADTDLLIDYVVDTTIKNTAKKG---GIYAKKAKMSRIDQLNISPNSATEIGGGICC 750
Db 634 TSSTSLLSNKKATARSLSNATSASNVPTGTFSMSSHSTSVITPGFSTSSA----- 684
QY 751 KESLELDALVSLVTENLVGKEGGLHAKTVNINLKSFGFSNNKANSSTGVTATTASA 810
Db 695 ---SLAINSTV-----VSSSLAGYSFSTPE-SSPTTSLTLVTSEA 719
QY 811 PAAAAASLQAA-----AAAAPSSPAT-----PTYSGVGGAIYGEKVTFSQ 851
Db 720 PSTVSSMTTSAPFINNSTSARPSSTASFITESSTSSVPLASGDVTSLLAAHNLTFSS 779
QY 852 CSGTCQFGNGAIDNNPQSSSLNV-----OGGAIYAK---TSL-----STGSSDAG 894
Db 780 APST---SSAQLVSKSTSSSLTVPTRIDRSGNSSTASRIATSLPNKTTFFVSSLSSTSAH 836
QY 895 TSYTFSGNSVSTGSKQTGQIAGGAIYSPVTLNCPATFSSNTASIAIATPKT---SSEGDSS 952
Db 837 ARNIFNSTVLATAKO-----IEFLTSTVNC-----SMTPTNYNTKTVIVSRETTA 882
QY 953 GNSIKDTIGGAI---AGTATILSGVRSFGNTADLGAAGIATLANANPPSATGSSONSIT 1008
Db 883 IGTVTSCEGCGTKNRKSTYTLITIDID-----ASTVITCEKEVTSYTGDEAHT 933
QY 1009 EKITLENGSTFIFERNQANKRGAIYSPSVSIKNNITFNQNTSTHDGSAIYFTKDA----- 1063
```

```
Db 934 TSTRISN-----FETS-----TFSE---SFKDMKTSQETKKAKPGSE 967
QY 1064 TIESLGSVLFTGNNVTATQASSATSGQNTNTANYGAAIFGDPGTQSSQTDAILTLASS 1123
Db 968 TVRSSSS---FVEKTSPTTKASPTSPSESkaa----- 997
QY 1124 GNITFSNNSLONNOGDPASKFCSTIAGYVKLSLOAAKGKTIISFFDCVHTSTKKTGST--- 1180
Db 998 GNTSVAINA-----SPSTSPSESQGTGSTSVEGAKSKSTKNSGEGVSTTKAKWTSTVAK 1050
QY 1181 -----QNVYETLDINKENSNPYGTIVFSSSELHENKSYIPQNAIHNGLTVLKE 1230
Db 1051 SSTESPGRGETTLEIIVVSQKSL--LTSQLSSSTE-KVNRSTTKPTAAIHGTSSSAKQ 1107
QY 1231 KTELHVVSFEQEKESKLIMEPDAVLNSNOIANGALAI 1267
Db 1108 STTYVTAKENTGASLNINMKAFVI-----GAIAL 1138
RESULT 5
OMPE_CHLTR STANDARD; PRT; 964 AA.
AC 084877;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE OUTER MEMBRANE PROTEIN E PRECURSOR.
GN MPPE OR CT869.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DJW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001360; AAC68467.1; -.
CC Outer membrane; Signal.
CC SIGNAL 19 964 PUTATIVE OUTER MEMBRANE PROTEIN E.
CC FT CHAIN 1 964
CC SEQUENCE 964 AA; 104703 MW; 1B998A7D2E571CE2 CRC64;
Query Match 4.4%; Score 394; DB 1; Length 964;
Best Local Similarity 23.5%; Pred. No. 3.4e-11;
Matches 269; Conservative 133; Mismatches 376; Indels 366; Gaps 60;
```

```
QY 653 PKSAPVST--ALSPSSST-----VSSLTLAASSQASPATSNKETQDPNADFDL 700
Db 29 PNVSPDPTKESLNKISLTGTHNLNCTYLDNLRYILAIQKTPNEGAAVT----- 79
QY 701 LIDYV--VDTT-----ISKN-TAKKGGGI-YAKKAKMSRIDQLNISENSAT-----EIG- 745
Db 80 ITDYLSEFDTKQEGYIFAKNLTPESGGAIGA-----SPNSPIVEIRDTIGP 126
QY 746 ----GGICCKSELELDALVLSLVTENLVG----KEGGGLHAKTVNINLKSFGFSNNKA 797
```

```
Db 127 VIFENNTCCR-----LFTWRNPYAADKIREGGA THAONLY INNHVDVVGFMKN--- 174
QY 798 NSSSTGVATTASAPAAAAAALQAAAAAPSPPATPTYSYGVGGAIYGEKVTFSQCSTGCTQ 857
Db 175 -----FSYVQGGAI-----STANTEV 190
QY 858 FSGNQA-----IDNPNQSQSLNVGGAIYAKTSLSTGSSDAGTGYFSGNSVTGKSGQTTC 913
Db 191 VSENOCSFLFMDNICIQTNTAGKGAIGAGTNSFESNCCDFFI-----NNA 238
QY 914 QIAGAIYSPVTTL---NCPATFSNNTASIAPTKTSSESSGNSIKDITIGGAIGACTAIT 970
Db 239 CCAGAGIFSPICSLTGNRGNIVFYNNRC--FKWVEFASSEASDGAIKVT-----TRID 290
QY 971 LSGVRSFGNTADLGAAIGCTLANANTPSATSGSONSITEKITLENGSFIFERN-QANKRG 1029
Db 291 VTG-----NRGRIFFSNDITKNYGG 310
QY 1030 AIYSPSVSIKGNNT-FNONTSTHDSGAIYFTKDATIESLGS-----VLTGNNVTATQA 1083
Db 311 AIYAPVTLVDNGPTYFINNIANNKGATYI--DGTSSKISADRHAIIFNENIVT----- 364
QY 1084 SSATSQNTNTANYGAAIFGDPCTTQSSQTD--ILTLASSGNITFSNNSLQNNQGDTP 1141
Db 365 -----NVTNAN-----GTSTAMPRRNALTIVASSSGEILLGAGSSON----- 402
QY 1142 ASKFCSIAGYKLSLQAARKTISFFDCVHTTKTGTQNVYETIDINKENSNPYTC 1201
Db 403 -----LIFYDPIEVS-----NAGVSVSFNKEADQ---TGS 429
QY 1202 IVF-----SSELHEN--KSYIPQNAIHLNGTLVLKEKTELHVSVFEQKESKLIMEPGA 1253
Db 430 VYFSGATVNSADFORNLOTKTPAPLTLSNGFLCIEDHAQLTVNRFTQ--TGGVVSGLNGCA 488
QY 1254 VLS-----NQNTANGALAINGLTIDLSSMGTPQAGEIFSPPELRIVATTSASGSGSVS 1307
Db 489 VUSCYKNGTGDSSASNASITLKHIGLNLSSI--LKSGAEI---PLLWVEPTNNNN----- 538
QY 1308 SSIPNPKRISAAPVSGSAATPTMSENKVFLTGDLTIDPNGNFYQNPMLGSLDLDVPLI 1367
Db 539 -----YTADTAATFSLSDVK-----LSLIDDYGN---SPYESTDLTHALS 575
QY 1368 KLP-----TNTSDVOYVDLTLSGDLFPQKGYNGTWTLD-----SNPQFG 1406
Db 576 SQPMLSISEASDNQSENIDFSLNPHYWGQGLTWGWAQTQDPEPASSATITDPQKA 635
QY 1407 KLOARWTFDTPYRWVYIPRDNH---FYANSILGSONSMITVVKQGLNNMLNNAFPDDIAY 1463
Db 636 NREHRTLLLTLPAGTVPSPKRSPLIANTLWG---NMLLATESL-----KNSAELTPSGH 688
QY 1464 NNFW---VSGVGTFLAQOQTPLSEEFYSYSRGTSVAIDAKPRQDFTLGAAFSKIYVKTKRA 1520
Db 689 -FPWGITGGGLMMYVQDPRENHPCGFHMSSGYSGAGMIAGOTHTFSL--KFSQTYTKLNE 745
QY 1521 IKKMNIYFHKGEYSYQASVYGGKFLYLLNKHQHWALPFLIOGVVSYG-HKHDDTFTLY 1579
Db 746 RYAKNNVSSK--NYSQ-----GEMLFSL---QEGFLLTKLV-GLYSYGDHNCHEFTYQG 794
QY 1580 PSIHENKGDWEDGLWDLRLISMDLKEPSKDSKRITV--YGELE-YSSIRKQKQFTED 1636
Db 795 ENLTSGTFRSQTMGG-----AYFFDLPMPKPGFTSHLTAPFLGALGIYSSL--SHFTEY 848
QY 1637 YDPRIHFD-DCAVRNLSLPGVCAVEGAIMNCLIMYNKLA---LAYMPSIYRNPNVCKYR 1691
Db 849 AYPRSESTKTPLINLVPIG--VKGSFMNAT---HRPQAWTVELAYQPVLYRQEPGIAAQ 903
QY 1692 VLSS 1695
Db 904 LLAS 907
RESULT 6
OMP_RICPR
```

DR EMBL; AJ235273; CAA15140.1; .  
KW Antigen; S-layer; Cell wall.  
FT CHAIN 1 1328 120 KDA SURFACE-EXPOSED PROTEIN.  
FT CHAIN 1329 1643 32 KDA BETA PEPTIDE.  
FT VARIAT 257 257 V -> A (IN STRAIN BREINL).  
FT VARIAT 1010 1010 Y -> D (IN STRAIN BREINL).  
FT VARIAT 1450 1450 A -> S (IN STRAIN BREINL).  
FT CONFLICT 178 179 AA -> VC (IN STRAIN BREINL).  
FT CONFLICT 191 201 TTEAPLTLGA -> INSRSSSYHLVS (IN REF. 1).  
FT CONFLICT 212 212 T -> I (IN REF. 1).  
FT CONFLICT 313 313 Q -> L (IN REF. 1).  
FT CONFLICT 1104 1104 T -> G (IN REF. 2).  
FT CONFLICT 1123 1123 T -> S (IN REF. 2).  
SQ SEQUENCE 1643 AA; 169854 MW; 735FDF392E6346CC CRC64;

Query Match 4.13; Score 371; DB 1; Length 1643;  
Best Local Similarity 21.06; Pred. No. 8.1e-10;  
Matches 373; Conservative 203; Mismatches 642; Indels 562; Gaps 87;

QY 2 KWLATAVEAAVLPSVSGFC-FPEPKELNFRSVETSSSTTETIG---EAGAEYIVSGN 57  
DB 10 KIISAGLVATAIVAGFSVGAAMQYRNTNAAATF-DGIGFQQAAGANIPVAPN 68  
QY 58 ASFTKFTNIPITDTPPTNSNS-----SSSGETASVSDSDSTTTTPDKGGGAFYNAHSG 114  
DB 69 SVITANANNPITPNCGLNSFLDTANDLAVTINEDT----- 107  
QY 115 VLSFWTSRGSTLSEIKMTGGGAIFS---OGELLFTDLTSLTTONNLSQLSGGAIF 171  
DB 108 TLGFTITNAQAQK-----FNTVAAGKILNITGGGITVQEAASNTINAQNALTKVHGGA 163  
QY 172 GGTISLSGITKATFSCNSAEVPAVPKPEPKAQTASSETSGSSS---SSGNDVSVPSS 228  
DB 164 NAN-DISGLGSIIFAAAPSVLENLNP---TQEPAPLTLGANSKIVNGNGTL----- 213  
QY 229 SRAEPAANLQSHFICATAPAAQTDTETSPSHKPGSGGAIYAKGDLTIADSOEVLFSI 288  
DB 214 -----NITNGFIQVSDNTFAGIKT-----INIDDCOGLMFN- 244  
QY 289 NKATKDGGAIFAEDKVSFENITSLKVOTNGAEKGGAIYAKGDLISQSSKOSLFSNSYSK 348  
DB 245 --STPDAA-----NTLNQVGGNTINFNGIDGTGKLVLSVSKNGAATFNFVTG 289  
QY 349 QGGGALYVEGGINFODLEEIRIKYKAGTFETKKTITPLSKAKAQAAGNADAWSSSPQS 408  
DB 290 TLGNL---KGI-----IEINTAA-----VAGKLISQGAANAVTGTONGAGRA 330  
QY 409 SGATTVSDGSSGSDSDTSETVPVTAKGGGLYTDKNLSITNITGIIIEIANNKATDVGG 468  
DB 331 AGFIVSDNGNAATIS-----QQVYA-KNMVIOQANAGGQVTFEHIVDVGL 375  
QY 469 GAYVKGTLTCENSHRLQFLKNSDKGGGIIGEDN-----ITLSNLTKTL 514  
DB 376 G-----GTFNFKTADSKVLIITENS-----FGSTNFGNLDFTQIVVPDTKILKGNFIGDVK 425  
QY 515 FOENTAK----EEGGGLFIKGTDKALMTGLDLSFCLINNTSEKHGGAFVTKESIQVTS 570  
DB 426 NNGWTAGVITFNANGALVSASTDNIAVTNI-----NAIEAEGAGVV---ELSGIHIA 475  
QY 571 DVEITPG---ITPVHGETVITG---NKSTGGNGGGVCTKRLALNSLQSIISGNSAAENG 626  
DB 476 ELRLNGGGSIFKLADGTVINGPVNQALMNNALAAAGSIQLDG---SAIITGDI---GNGG 530  
QY 627 AHTCPDSEPTADTAEOAPAAASATSTPKSAPVSTALSTPSSSTVSSLTLLAASSQASPAT 686  
DB 531 VNAALQHITTLANDASKILALDGANII---GANVGGAIHFQANGGTIKET----- 576  
QY 687 SNKETQDPNADTDLIDYVVDITISK-----NTAKKGGIYAKKAKMSRIDOL 734  
DB 577 -----NTQNNIVNFNFDLITDKTGCVVDASSLTNQTUTINGSIGTVVANTKTIAQL 628  
QY 735 NISENSATEIGGGICCKESLELDALVLSVTENLVKREGGLHAKTVNISN----- 785

DB 629 NIG-SSKTIILNAGDVAINELVIENNGSVQLNHNHY-----LITKTIANAANQOLIIVAAD 681  
QY 786 -----LKSG-----FSPSNKANSSS-----TGV---ATTASAPAAAAASL 818  
DB 682 PLNTNTTTLADGTNLGSAENPLSTTHEFATKAANADSILNVGKGVNLYANNITNDANVGS 741  
QY 819 QAAAAAAPSPATPTYGVVGGAIYGEKVTFSCSGTQCFSGNQAIIDNPSSOSL---NV 875  
DB 742 HF-----RSGGTSIVSGTVGGQ-QGHKL-----NNLILDNGTIVKFLGDTTF 782  
QY 876 QGGA-IYAKTSLSGSSDAGTSYIFSGNSVSTGKSQTTGQIAGGAIYSPVTLN----- 928  
DB 783 NGGKIEGKSLIQI-SNNYTTDHVESADNTGTLEFVNTDPI-----TVTLNKGAGYF 833  
QY 929 -----CPATESNN-----TASIATPKTSESGSSGNS-----TKDT 959  
DB 834 GVLQVITISGPNIVFNEIGNVGIVHGAANSISFENASLGTSFLPSGTPDLVLTIKST 893  
QY 960 IGGAIAGTA-----ITLGSVSRFSGNTADLGAAGTLANANTPSATSGSONSITEKI- 1011  
DB 894 VGN---GTVDNFNAPIVVVSGIDSMINN---GOIIGD--KKNIIALSGLSDNSITVNAN 944  
QY 1012 TLENGSIFERNQA-----NKGAIYSPSVSIKGNITFNQNTSTHDGSAIYFTKD 1062  
DB 945 TLYSGIRTTKNOGTVTLSGGMPNPGTIY--GLGLENGSPKLQVFTTDD----- 993  
QY 1063 ATIESLGSVLETTGNVTVATQASSATSGONTNTANYGAI-----FGDPGTQTS 1110  
DB 994 --YNNLSGII--ANNVTINDYVTLTGTGIAGT-DFDAKITLGSVNGNANVRFPD--STFS 1046  
QY 1111 SQTDAIITLASSGNITFSNNSLONNOG--DTPASKFCSIAGYVKLSLOAAKGTISFFD 1168  
DB 1047 DPRSMIVATQANKTVTYLGNALVNSIGSLDTPVA----- 1081  
QY 1169 CVHFTSKTKGTQ-----NVY-ETLDINKEE-----NSNPYT--GTIVFSSELHENKSYI 1215  
DB 1082 ---SVRFTGDSGAGLQGNISQNIQDFGTNLTLNSNVILGGTGAINGEI-----DLL 1133  
QY 1216 PQNALHNGTILVLEKTE---LHVVSFEQKEGSKLIMEPAGV-----LSNQNIAN 1262  
DB 1134 TNNLIFANGTSTWGDNTSISTTLNVSS--GNIGQVIAEDAQVNAATTTGTTTIKIQDNAN 1191  
QY 1263 -----GALAINGLTIDLSSMGTPQ---AGEIFSPPEL----- 1291  
DB 1192 ANFSCTQATYLIQGGARENG-----TLGAPNFAVTSNIEVKYELIRDSNQDYVLFTN 1245  
QY 1292 RIVATTSASGGVSSSIPNPKRISAAPVSGSAATPTPMSENKVFUTGDLTLDIPNGN 1351  
DB 1246 DVLNVVTTAVGNSAIAA-PGVSONISRCLES-----TNTAAYNNMLIAKD----- 1290  
QY 1352 FYQNPMLGSLDVLPLIKLPTNTS---DVQVVDLTLSGDLFPQKGYAGCTWLDNSNPQTGK 1407  
DB 1291 -----PSDVATFVGAIAATDTSAAVTTVNLNDTKQDQLLSNR--LGTLYLSNAETSD 1341  
QY 1408 LQARWTFETYRRWVIIPRDNHFYANSILG---SONSMIVVKOGLINNMNNARFDDIAYNN 1465  
DB 1342 V-----AGSATGAVSSGDEAEVSYGVWAKPYNYNAEQDK--- 1376  
QY 1466 FWVSGVGTFFLAQOGTPLSEEFYSYRGTSTVAIDAKPRODFILGAAFSKIVGKTKAIIKMH 1525  
DB 1377 ---GGIAGYKAK-----TTGVVVGDLTSLASNLMIAGAA---IGITKTDIK-H 1416  
QY 1526 NYFHKGSEYSQA---SVYGGKFL-----YFLNKK 1552  
DB 1417 QDYKKGDKTDINGLSFLSYGSQLVKNFFAQNIAFTLNK 1456

RESULT 7  
OMP\_RICRI  
ID OMPBL\_RICRI STANDARD; PRY: 1654 AA.  
AC Q53047;  
DT 30-MAY-2000 (Rel. 39, Created)





```
QY 1244 GSKLIMEPGLVSLNONTANGALINGLTLIDLSMGT-----PQAGEIFSPPE-LRIVAT 1296
D 1152 GNNTSIETTLFLANGIGH-IVILEGAQVNTTTTGTITIKVQVDNANANFSGTQTYTLIQ 1210
QY 1297 TSSASGSGVSSSIPTNPKR-----ISNAVPSGSAATPTMSENKVF LTGLD TLIDP N 1351
D 1211 GARFNGTLGSPNFAVTGSNRFVNTSLIRAAHQDYVITRTNNAENV--TNDIA-----NSP 1264
QY 1352 FYQNPMLGSDLDVPLIKLPTNTSDVQVVDLTLSGLDFLPQKGYMGTWTLDSPQTKLOAR 1411
D 1265 FGGAP--GVDONVITFVNATTAAYN--NLLAKNSANSANFVGAIVTDSAAITNVOLD 1320
QY 1412 WFTDYRRWVYIPRDNHYANSILGSQNS--MIVVKQGLINML--NNAREDDIAYNNFW 1467
D 1321 LAKD-----IQAQLGNRLGALRGITPETABMAGPEAGISAAVAAGDEAIDNVAI-GIW 1374
QY 1468 VSGVGTFLAQQGTPLSEFYSYSGTSAIDAKPRQDFILGAASFVKTKAIKKHNY 1527
D 1375 AKPYTDAHQSKKGLAGYKAKTTGVVIGLDTLANDLNLMIGAA-----IGITKTDIK-HQD 1429
QY 1528 FHKGSE---YSYQASVYGGKFLYELLNKHQGWALPFLIQ--VVSYGHKHKHDTTTLVPSI 1582
D 1430 YKKGDKTDVNGFSFSLGAQQL-----VKNFAQGSIAFSLNQVKNKSQRYFEDA 1479
QY 1583 H-----ERNKGDWEDGLWADLRISMDLKEPSKDSKRITVYGELEY-----S 1625
D 1480 NGNMSKQIAAGHYDNWTFGGNLTVGYDY---NAMQGVLTVPMAGLSVLKSSDENYKGTG 1536
QY 1626 SIROKQTEIDYDRPHDDCAYNRLSLPVGCAVEGAINMNCILMYNKLALAYMPSIY 1682
D 1537 TVANKQVNS-----KFSQD-----RTDLIVGAKVAGSTMNIT-----DLAVYPEVH 1576

RESULT 8
OMPF_CHLTR
ID OMPF_CHLTR STANDARD; PRT: 1034 AA.
AC P38008; O84878;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE OUTER MEMBRANE PROTEIN F PRECURSOR.
GN PMPF OR CT870.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/OW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.
RA "Genome sequence of an obligate intracellular pathogen of humans:
RA Chlamydia trachomatis.";
RT Science 282:754-759(1998).
RN [2]
RP SEQUENCE OF 26-35.
RC STRAIN=L2/434/BU;
RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA Comanducci M., Christianen G., Birkelund S., Vitreou E., Ratti G.,
RA Pallini V.;
RA Submitted (SEP-1994) to the SWISS-PROT data bank.
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC - EMBL; AE001360; AAC68468.1; -.
```

```
KW Outer membrane; Signal.
FT SIGNAL 1 25
FT CHAIN 26 1034 PUTATIVE OUTER MEMBRANE PROTEIN F.
FT CONFLICT 34 35 RR -> RH (IN REF. 2).
SQ SEQUENCE 1034 AA; 112392 MW; 445FF4C35D463AE7 CRC64;

Query Match 4.0%; Score 359.5; DB 1; Length 1034;
Best Local Similarity 21.9%; Pred. No. 1.5e-09;
Matches 244; Conservative 158; Mismatches 441; Indels 269; Gaps 50;

QY 778 AKTVNISNLKSGFS-----FSNNKANSSTGVATTASAPAAAAAALQAAAAAPSSPATP 832
D 55 AQIVLSNLSNGTGACTISGNTQTQIFSNVNTADSGA----- 95
QY 833 TYSGVVGAIYGEKVTFSQCSGTQCFQFS-GNOAIDNNPSQSSLVNVOGGAIYAKTSLISGSS 891
D 96 -FDMVTTSTASDNANILFCNNYCTHNKGGGAIRGPGPIREFLNQDVLFYNNIS----- 148
QY 892 DAGTSYIFSGNSVSTGKSTQTTQIAGGAIYSPTVTL--NCPATFSNNTASITATPKTSED 949
D 149 -AGAKYVGTGDHNEKR-----GGALYATTITLTGNRTLAFINNMGGDCGGAISAD- 198
QY 950 GSSNSIKDTIGGAIAGTAITLGSVSRFSGNATDLGAAIGTLANANTPSATSGSONSITE 1009
D 199 --TQISITDTVKGILFENNHTLNHPYTAENMARGGAI-----CSRRLDLCISNNS--- 248
QY 1010 KITLENGSFIFERNOANKRGAIYSPSVSIKGN-----NITFNQNTSTHDSAIYF 1059
D 249 -----GPVFNYNGOGKGAISATRCVIDNNKERIIFSNNSILGWSQSSASNGAIQT 302
QY 1060 TKDATI--ESLGSVLFTGNVTATQASSATS-----GQN-----TNTANYGAAIFGDPG 1106
D 303 TQGTFLRNKGSIIYFDSN--TATHAGGAINCGYDIRNGPVPYFLNNSAANGAF----- 355
QY 1107 TTQSSQTDAILTLASSGNITFSNNSL-----ONNQDGTPAKFCSCIAGY 1151
D 356 -NLSKPRSATNYIHTGTGDIVFNNNVFTLDGNLLGKRKLPHINNNEITPYT----- 406
QY 1152 VKLSLOAAKGTISFFDCVHTSTKTKGSTQN-----VYETLDINKEENSNYPTGTVFS- 1205
D 407 --LSGAKKDTRIYFYDILFQWERVKENTSNPPSPTRNTITVANPE---TFSGAVVFSY 461
QY 1206 -----SELHNKSYIPO--NAILHNGTLVLKTEKTELVH--VSFEQEGSKLMPEGAV 1254
D 462 NQMSDITRLMGKEHNYIKEAPTTLKFTCLAIEDDAEIEINIFTQNTPTSLALGSCAT 521
QY 1255 LSNQNIANGALAINGLTIDLSSMGTPOAGEIFSPPELRIVATTSASGGVSSSIPTNP 1314
D 522 LTVGK--HGKLNITNLGVILPII----LKEGKSPPCIRVNPQDMTQNTGTGOTPS----- 570
QY 1315 KRISAAVPSGSAATPTMSENKVF LTGLD TLIDP N-GNFYQNPMLGSDLDVPLIKLPTNT 1373
D 571 -----STSSISITP-----MIIFNGRLSIDENYESVYDSMDLSRGKAEQILSIIETT 617
QY 1374 SDQVQ---YDLTSLGDLF--PQKGYMGTW-----TLDSPNQ-----TGKLOARW 1412
D 618 NDGOLDSNWQSSLNTSLLSPHYCYQGLWTPNWTITVITLNNSSAPTSAISAEQKK 677
QY 1413 TFDYQ----RRWVYIPRDNHYANSILGSQNS--MIVVKQGLIN--MLNNAREDD 1460
D 678 TSETFTPSNTTASIPNIKASAGSGSASNGEVTITKHTLVVNWAPGVIVDPPIRGD 737
QY 1461 IAYNNEWVSGVGTFLAQQGTPLSEFYSYSGTSAIDAKPRQDFILGAASFVKTKA 1520
D 738 LIANSLVHSGRNWMTGURS--LLPDNSWFA-----LOGAATLTETKQKR 780
QY 1521 IKKMHNHY--FHKGSEYSYQAS--VYGGKFL-----YFL--LNKQ 1553
D 781 L-SYHGYSASKGYTVSSQASGAHGKFLLSFSOSSDKMKEKETTNNRLSSRYLSALCFE 839
QY 1554 HGWALPFIQGVSYGHKHKHDTTTLVPSIHERNKGDWEDGLWADLRISMDLKEPSKDS 1613
```









```
QY 1088 -SQONTNTANTGAALFGDPGTTQSSQTDAILTLASSGNITPNNSL-----Q 1134
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1328 SSKSSLSLSPVSSSILMSQFSSSSSSSSSLASLSISPVDVTVSVLPQPTTSIATLCT 1387
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1135 NNQGTDPKPC--SIAGYVKLSLQAAGKTIISFPDCVHTSTTKTKGSGNQNVETLDINKE 1192
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1388 DSQCQOEVTICGNSCDDVTSATTPPSTVDTMTCTGSECQKTSS-----SCDGYSC 1442
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1193 ENSNPTCTIVES-----SELHENKSYIPQNAIHLNGTLVLKKEKTELHVVSFE 1240
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1443 KVSEYTKSATISAGSGGCGQASATSEL--NSQYVTMTSVI--TPSAITTSVEVH----- 1494
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1241 QREGSKLIMEPCAVLSNONGIANGLAINGLITDLSSMGTPOAGEIFSPPELRIVATTSSA 1300
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1495 STESISITTVKPYVYTSSTNGEL-----ITITSSQ 1527
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1301 SGGSGVSSIPNPKRIISAAVPSGSAATPTPMSENKVFELTGDLT 1344
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1528 T----VIPSVTTIIRTKVAITSAPKPTTTTIVVEQRLSSSGIAT 1567
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 12

```
ICEN_XANCT
ID ICEN_XANCT STANDARD; PRT; 1567 AA.
AC P18127;
DC 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ICE NUCLEATION PROTEIN.
GN INAX.
OS Xanthomonas campestris (pv. translucens).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas
OX NCBI_TaxID=343;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91080859; PubMed=2259339;
RX Zhao J., Orser C.S.;
RA "Conserved repetition in the ice nucleation gene inax from
RT Xanthomonas campestris pv. translucens.";
RL Mol. Gen. Genet. 223:163-166(1990).
CC -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CC CRYSTALLIZATION IN SUPERCOOLED WATER.
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
CC -!- DOMAIN: CONTAINS 133 IMPERFECT REPEATS OF THE CONSENSUS
CC OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52970; CAA37140.1; -.
CC PIR; S11672; S11672.
CC HSP; P06620; IINA.
CC InterPro; IPR000258; -.
CC Pfam; PF00818; Ice_nucleation; 81.
CC PRINTS; PR00327; ICENUCLEATFN.
CC PROSITE; PS00314; ICE_NUCLEATION; 57.
CC Ice nucleation: Repeat; Outer membrane.
CC SEQUENCE 1567 AA; 152548 MW; C8B451D959ECAD63 CRC64;
```

```
Query Match 3.8%; Score 341.5; DB 1; Length 1567;
Best Local Similarity 22.0%; Pred. No. 1.8e-08;
Matches 276; Conservative 185; Mismatches 537; Indels 259; Gaps 56;
QY 67 PTTDTTTTNNSSSSSSSETASVSDSDSTTTTPDKGGGAFYNAHSGVLSFMTSGTEG 126
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 PSTQATPSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQSTQ 207
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 SLTSL--EIKMTGGGALFSGGELLFTDL-----TSLTIONNLSOLSGAIFGGSTISL 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 S-TLGAQDSRLVAGYGSTETAGD--HSDLIAGYSTGTAGSDSSILAG-----YGSTQTA 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 SGITKATFSCNSAEYAPAPVKPTEPKAQTASSETS-----GSSSSSGNDS--VSSPSSR 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 AGRSLTAGYGS-----TQTAQEGSLTSCYGGSTATSGSDSAVISCYGSTQ 306
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 AEPAAANLQSHFICATYATPAATDTETSTPHSHKPGSGGA-----IYAKGDLTITADSOEVL- 285
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 TAGSESSLTAGY--GSTQTAARKGSDITAGYGSTGTAGSDSALIAGYSTGTAGSESSLT 363
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 --FSINKATKOGGAI--FAEKDVSFENITSL---KVQTNCAEEKGGAIIYAKGDLISIQS 336
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 AGYGSTQTAARKGSDVDTAGYGSTGTAGADSTLIAGYGSTGTAGGESSLTAGYGSTQTARQG 423
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 SKQSLFNSNYKQGGGALYVEGGINFQDL-----EERIKNYKAGT 377
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 SDITAGYGSTGTAGADSTLIAGYGSTQTSQSDSSLTAGYGSTQTAARKGSDITAGYGSTGT 483
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 FETKKIITLPSLKAQASAGNAD-----AMASSPQSGGATTVSDSDSSS---- 422
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 484 AGSDSSLIAGYGSTGTAGSESSLTAGYGSTQTAQDSSLTGTGYGST--STAGHDSLIAG 541
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 423 -GSDSDTSETVPVTAKGGLYLT-DKNLSLTNTITGIIETANNKATDVGGAIVKGLTTCEN 480
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 YGSTGTAGYDSTLIAGYGSTQTAQDSSLTGTGYGSTGTAGADSTLIAG---YGSTQTAGS 598
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 481 SHRL-----QFLKNSSDKQGGGIYGEDNITLSNLTQKTLFQENTAKKEGGGLFIKGTID 533
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 DSSLTAGYGSTQTAREGSDVTAG--YGS-----TG-TAGADSTLIAGYGSTQTSQSD 647
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 534 KALTWTGLDLSFCLINNTSEKHGGGAFVKEISQTVT--SDVETIPGTPVHGETVITGNK 591
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 648 SSLT-AGYGS---TQTARK--GSDVTAGYGSTGTAGADSTLIAG---YGSTQTSQSD 695
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 592 S--TGGNGGGVCTKR-----LALSNLOSISISGNSAAENGGAHTCPDSPPTADT 639
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 696 SSLTAGYGSTQTAARKGSDVDTAGYGSTGTAGADSTLIAGYGSTQTSQSDSSLTAGYGSTQT 755
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 640 AEPAAAAAATSTPKSAPVSTAL-----STPSSSTVSSLTLLAASSQASPATSN-----K 689
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 756 ARKGSVDVTAGYGSTGTAGADSTLIAGYGSTQTSQSDSSLTAGYGSTQTAARKGSDITAGY 815
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 690 ETQDPNADTDLIDYVDVTI-----SKWTAKKGGGIYA--KKAKMSRIDQLNI 736
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 816 STGTAGADSTLIAGYGSTQTSQSDSSLTAGYGSTQTAAREGSDVTAGYGSTGTAGADSTLI 875
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 737 SENSATEIGGGICCKESLELDALVLSVTENLVGREGGLHAKTVNLSNLSKGSFSSNNK 796
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 876 SCYGSTQTAGS-----DSSLTAGYGSTQTARKG-----SDVTAGY-G-STGT 915
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 797 ANSSST-----GVATTASAPAAAAAALQAAAAAPSSPATPTYSGVVGGAIIYGEKVFESQ 851
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 916 AGADSTLIAGYGSTQTSQSDSSLTAGYGSTQTAARKGSDMTAGY-GSTGTA--GADSTLIA 972
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 852 CSGTCQFSGNQADNNPSSQSLNVGGAIYAKTISLIGSSDAGTSYIFSGNSVSTGKSTQ 911
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 973 GYGSTQTSQ-----SDSSLT---AGYGSTQTAAREGSDVTAGY--G-STGTAGADST 1017
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 912 TGOIAGGAIYSPTVTLNCPATFSNNASTATPKTSE---DGSSGNSIKDTIGGAIGACT 967
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1018 L--IAG---YGSTQTAGSDSSLTAGYGSTQTAARQSDVTAGYGSTGTAGADS-----T 1065
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 968 AITLSGVSRSFGNTADLGAAIGTLANANTPS-----ATSGSONSITEKITLENGSF 1018
```

```
Db 1066 LIAGYSTQTAGSDSLTAGYGTQARQSGDITAGYGTGTAGADSSLIAGY----GST 1121
QY 1019 IPERNOAKRGALYSVSIKGNITF---NONTSTHDSGAIYFTKDATIESLGSVLEF- 1074
Db 1122 QTAGYDSNLTAGYGTQATARESSSLTAGYGTSTAGHDSLIAGYGTQTAGYNSILTTG 1181
QY 1075 -GNVTVTAQASNTSGO-NTNTANYGAAIFGDPGTGTOSSQTDAILTLASSGNITFSNNS 1132
Db 1182 YGSTQTAQESSSLTAGYGTSTAGYDSTLTAGYGTQTAGYKSTLTAGYSGNSTAGHSS 1241
QY 1133 LQNNQDTPASRCSAGYVYKLSLQAAKGTISFFDCVHTSTKTKTSTQNYETLDI 1189
Db 1242 LIAGYST-----QIAGY-ESTLTAGYSSSL-----TQOSSLTAGYSGSTEI 1283

RESULT 13
OMPBLRICYV
ID OMPBLRICYV STANDARD; PRT; 1645 AA.
AC P96989;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OUTER MEMBRANE PROTEIN B PRECURSOR (168 KDA SURFACE-LAYER PROTEIN)
DE (SURFACE PROTEIN ANTIGEN) (CELL SURFACE ANTIGEN) (SCA5) (ROMP B)
DE (CONTAINS: 120 KDA SURFACE-EXPOSED PROTEIN (SURFACE PROTEIN ANTIGEN)
DE (120 KDA OUTER MEMBRANE PROTEIN OMPB); 32 KDA BETA PEPTIDE].
GN SLP.
OS Rickettsia typhi.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxId=785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WILMINGTON;
RX MEDLINE=94040787; PubMed=8224896;
RA Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;
RT "Cloning and sequence analysis of the gene encoding the crystalline
RT surface layer protein of Rickettsia typhi.";
RL Gene 133:129-133(1993).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN=WILMINGTON;
RX MEDLINE=92114896; PubMed=1370573;
RA Ching W.M., Carl M., Dasch G.A.;
RT "Mapping of monoclonal antibody binding sites on CNBR fragments of
RT the S-layer protein antigens of Rickettsia typhi and Rickettsia
RT prowazekii.";
RL Mol. Immunol. 29:95-105(1992).
RN [3]
RP IDENTIFICATION OF CLEAVAGE SITE.
RX MEDLINE=92104668; PubMed=1729180;
RA Hackstadt T., Messer R., Cieplak W., Peacock M.G.;
RT "Evidence for proteolytic cleavage of the 120-kilodalton outer
RT membrane protein of rickettsiae: Identification of an avirulent
RT mutant deficient in processing.";
RL Infect. Immun. 60:159-165(1992).
CC -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L04661; AAB48987.1; -
```

```
KW Antigen; S-layer; Transmembrane; Cell wall.
FT CHAIN 1 1353 120 KDA SURFACE-EXPOSED PROTEIN.
FT CHAIN 1354 1645 32 KDA BETA PEPTIDE.
FT TRANSMEM 1415 1429 MEMBRANE ANCHOR (POTENTIAL).
FT CONFLICT 657 657 H -> N (IN REF. 2).
FT CONFLICT 842 842 V -> I (IN REF. 2).
FT CONFLICT 1071 1071 G -> A (IN REF. 2).
FT CONFLICT 1306 1306 G -> S (IN REF. 2).
SQ SEQUENCE 1645 AA; 169698 MW; 0CB5641C7EB185EE CRC64;

Query Match 3.8%; Score 339.5; DB 1; Length 1645;
Best Local Similarity 19.0%; Pred. No. 2.3e-08;
Matches 377; Conservative 245; Mismatches 709; Indels 655; Gaps 92;

QY 2 KWLSTAVFAAVLPSVSGFC-FPEKELNFSRVETSSST-----FTETIGEAGAEIVS 55
Db 10 KIISAGLVASTATIVAGFCVGVAMGVQVNRNTNAAATVDGAGFDT--GAGVNLPA 67
QY 56 GNASFTKFTNPTTDTPTTNSNS-SSSGETASVSDSDSTTTTTPDPKGGGAFYNAHSG 114
Db 68 TNSVITANSNNAITFTPNGLNSLFLDTANTLAVTINENTTL-----G 111
QY 115 VLSFMTSRTEGSLTL---SEIKWTGEGGAIFSGOGLLFTDLTSLTIQNLSQSGAIF 171
Db 112 FVTNVTKQGNFFFTIGAGKSLTITGHG--ITAQ-----QAATTKSAQNVYVKVNAFAAI 164
QY 172 GGSTISLSGITTKATKSCNSAEVPAPVKPTEPKAQTASETSGSSSSGNDVSVSPSSRA 231
Db 165 NDN--DLSGVGSIDFTAAPSVLEFNINPTTQEA----- 196
QY 232 EPAANLQSHFICATATPAAQTDTETSTPSHKPGSGGAIYAKGDLTIADSOEVLFSINKA 291
Db 197 -----PLTGDNAKIVGANGI 213
QY 292 TKDGGAIPEAKVDVFNITSLKVOTNGAEKGGAIYAKGDLTSSQSKQSLFNSYISQGG 351
Db 214 LNTNGFVKVSDKTFAGIKTINI-----GD-----NQGLMFT--TPDAA 251
QY 352 GALLYVEGGINFQDLEIRIKYKAGTFETKKITLPSLKAQASAGNADAWASSPQSSGA 411
Db 252 NALNLOGGN-----TINFN--GRDGTGLVLVSKNGNATEFN----- 287
QY 412 TTVSDSGDSSGSDSDTSETVPVTAKGGGLYTDKNLSITNITGIE-----IANN 461
Db 288 -----VTGSLGG-----NLKGVIERDPTTAAAGKLIANG 315
QY 462 KA-----TDVGGGAVYKGLTCENSHRLOFLKNSDKGGGIYGEDNITLS-NLTGKTL 514
Db 316 GAANAVIGTDNGAGRAAGFIVSDVNGNAATI-----SCQVYAKDIVIQSANAGQVT 367
QY 515 FOENTAKEEGGLFIKGTDKALMTGLDSECLINNTSEKHGGGAFVTKETISQTYTSIVET 574
Db 368 FEHLVDVLOGGKTNFKTADSKVIITENASF-----GSTUFGN 404
QY 575 IPGITPVHGETVITGN---KSTGGNGGGVCT-----KRLALSNLQSIISIG 617
Db 405 LAVQIVVPNNKILTGNFIGDAKNNGNTAGVITFNANGTLVSGNTDPNIVVYVNIKAIEVG 464
QY 618 NSAEANGG--GAH-----TCPDSFPTAD-----TAEOPAAASAATSPKSPVPS 659
Db 465 AGIVOLSGIHGAELRLGNAGSIFKLADGTVINGVQNPLVNNNALAAGSIQDGSIAIT 524
QY 660 -----TALSTPSSS--TVSSLTLAASS-----QASPATSNKETQDPNAD 697
Db 525 GDTGNGAVNNAALODITLANDASKILTLISGANIIGANAGAIHFQANGGT-----IQLTSTQ 580
QY 698 TDLIDYVVDVT-----ISKNTAKGGGIYAKKAKMSRIDQLNISENSATEIG 745
Db 581 NNILVDFDLDTVDQTCVVDASSLTNNQTLTNGSICITICANTKTLCGRFNVGSSKTLNA 640
QY 746 GGTCCCKE-SLELDALVSLSVTENLVKGGGGLHAKTVNISNL----- 786
```

```
Db 641 GDVAINELVMDGSGVHLTHNTYLI-----TKTINAANQKIIVAADPINTDTALAD 692
Qy 787 -----KSGFSSNNKANSSS-----TGV-----ATTASAPAAAAAASLQAAAAAPSS 828
Db 693 GTNLGASPLNIIHFKAAKNGDSILHIGKGVNLYANNITTTDANVGLHP-----RS 746
Qy 829 PATPTYSGVVGAIYGEKVTFSQCSTQFCQSGNOAIDNPS---QSSLNVOGGA-IYAKT 884
Db 747 GGTISVGTGGO-QGLKL-----NNLILDNCTTVKFLGDTFNGGTKEGKS 793
Qy 885 SLUSIGSSDAGTSYIFSGNSVSTGKSTQGTQIAGGAIYSPVTVLCNPATF-----933
Db 794 ILQI--SSNYITDHIESADNTGTLEFVNTDPI-----TVELNRQGAIFYGLVKQVMVSG 844
Qy 934 -----SNNTA-SIATPKTSSDGGSGNS-----IKDTIGGAIAGTA-- 968
Db 845 PGNIAPNEIGNVGHAIADVDSISFENASLGASLFLSTGTPDLVLTIKSTVGN---GTVDN 901
Qy 969 -----ITLSGVSRFSGNTADLGAATGLANANTPSATSGSONSIT-EKITLENGSIPIER 1022
Db 902 FNAPILVSGIDSMINN---GOVIGD--QKNIIALSILGSONSITVNSNTLYAGIRTTKT 955
Qy 1023 NOA-----NKGAIYSPSVSTKGNITPNQNTSTHDSGAIYFTKDATIESLGSLVF 1073
Db 956 NQGTVTLSGGIPNPNPTIY--GLGLENGDPKLKQVTFTTD-----YNNLGSIIA 1002
Qy 1074 TGNVNTATOASSATSGONTANTANYCAAFGD-----PGTTOSSOTDAILTLASSG 1124
Db 1003 T--NTVINDVTLTGGTAGTDFDKITLGSINGNANVKFVDRTFSHTPSMIVSTKANOG 1060
Qy 1125 NITFSNNLSQNNQ--DTPASFCSCIAGVVKLSLQAAKGTISFPDCVHTSTKTKGQSTQN 1182
Db 1061 TVTYLGNALVGNIGSDIPVASVRFTGNDSGVGLGN-----IHSONIDFGT--- 1107
Qy 1183 VYETLDIINKEENSYPYGTIVFSPSELHENKSYIPONAILHNGTLVLEKTELUH-VVSFEQ 1241
Db 1108 -YNLILNSDVLG--GGTTAINGEI---DLTNNLIFANGTSTWGNNTSLSTTLNVSN 1160
Qy 1242 KEGSKLIMEPGA-----VLSNON-----IANGALATNGLTIDLSSWG 1278
Db 1161 GNVGQIVIAEGQVWATTTGTTTIKIQDNANANFSGTQTYTILIOGGARFNG-----TLG 1214
Qy 1279 TPQ-----AGEIFSPPEL-----RIVATTSASGGSGVSSSIPTNPKRISAA 1320
Db 1215 APNEDVTGNPIPVKVELIRANDANQDYVLRTNDVLNVVITAVGNSAIAANA-PGVHQNIAIC 1273
Qy 1321 VPSGSAATPTPMSEKNVFLTGDLTLIDPNGNYQNPMLGSLDVLPLKLPNTW----SDV 1376
Db 1274 LES-----TDTAAYNNMLLAKD-----SSDVATFIGAATDTGAAVATV 1312
Qy 1377 QYVDLTLSGDLFPQGYMGCTWFLDSNPOTGKLOARWTFDTYRRWYIPRDNHFNANSILG 1436
Db 1313 NLNDQTKQDOLLGNR--LGLARYLSNSETA-----DVG 1344
Qy 1437 SONSMVVKQGLINNLNNAAREDDIAYNFWVSGVGTFLAQOQTPLSEBFYSYRGTSTA 1496
Db 1345 SET-----GAVSS--GDEAIDQVSY-GVWAKPYNTAEQDKGLAGYKAKTAGVVVG 1394
Qy 1497 IDAKPRQDFILGAATSKIVGKTKALKKHNPYHKSEYSYQA---SVYGGKFLYELLANKQ 1553
Db 1395 LPTLANDNLMIQAA-----IGITKTDIK-HODYKKDKDITIKGLSFLSYCAQOL----- 1442
Qy 1554 HGWALPFLIQ--VVSYGHIKHDTT-LVPSIHERNK---GDWEDLGNLADLRISMDLK 1606
Db 1443 ---VKNFAQSGAIFTLNKKVSKSORYFFDANGKKNKQIAGNYDNITITGGNLMFGYDY- 1498
Qy 1607 EPSKDSKRITVYGELEYSSIRKQFTEIDYDPRH-----FDDCAYRNLSLPVGCAREG 1660
Db 1499 --NALQGVLTPTMAGLSYLKSSNENYKGTGTTVANKRIHRSKFS---RIDLIIVGAKVTG 1552
Qy 1661 AINMNCN-ILMYNKLALAYMPSIYR-----NNPVCKYRVLSSNEAGQVIGCVPTTSARA 1713
Db 1553 SAMNINDIVY-----PEIHSFVHVHKVNGKLSKAOSMLDQGTAPFI-SQPDRT-AKT 1602
```

```
Qy 1714 EYSTOL 1719
Db 1603 SYNIGL 1608

RESULT 14
CBPA_CLOCL STANDARD; PRT; 1848 AA.
ID CBPA_CLOCL
AC P38058;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CELLULOSE BINDING PROTEIN A PRECURSOR.
GN CBPA.
OS Clostridium cellulovorans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
CC Clostridium.
OX NCBI_TaxID=1493;
RN [1]
RP SEQUENCE FROM N.A. PubMed=1565642;
RX MEDLINE=92228810;
RA Shoseyov O., Takagi M., Goldstein M.A., Doi R.H.;
RT "Primary sequence analysis of Clostridium cellulovorans cellulose
binding protein A.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3483-3487(1992).
CC -!- FUNCTION: BINDS TO CELLULOSE FIBERS AND COORDINATES CELLULOSE
ENZYMES.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- PTM: CONTAINS ABOUT 10% OF CARBOHYDRATES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC
CC EMBL; M73817; AAA23218.1; -.
CC PIR; A44140; A44140.
CC HSSP; Q06851; INBC.
CC InterPro; IPR001956; -.
CC InterPro; IPR002102; -.
CC Pfam; PF00942; CBD_3.1.
CC Pfam; PF00963; Cohesin; 9.
CC Cellulose degradation; Cell wall; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 28
FT CHAIN 29 1848
FT DOMAIN 29 189
FT SEQUENCE 1848 AA; 189152 MW; 85FA6CE6F771AF1A CRC64;
SQ
```

Query Match 3.7%; Score 334.5; DB 1; Length 1848;  
Best Local Similarity 20.4%; Pred. No. 4.7e-08;  
Matches 380; Conservative 265; Mismatches 771; Indels 443; Gaps 80;

```
Qy 4 LSATAVFAVLPSVSGFCPEPKELNF-----SRVTSSSTTFTETIGAGAIYVSGNAS 59
Db 11 LALMVFALVLPISIPALAATSSMEFYNSNKAQNTSITPIKITNTSDSL-----NLN 66
Qy 60 FTKETNIPDTDTTPT-----NSNSSSSGSETASVSDSTTTTPDPKGGGA 107
Db 67 DVKRYRYTSDGTQCTQCTWCDHAGALLGNSYVDNTSKVTANFVKETASPTSYD----- 120
Qy 108 FYNAHSGVLSPFMRSGTEGLTSLSEIKMTGEGGAIFSGCELFTDLTSLTIONNLS---- 163
Db 121 -----TYVEFGFASGRATLK-----KGQFITIQGRITKSDWSNVTQNDYSFSDAS 165
Qy 164 -----QLSGGAIFG---GSTI---SLSGITKATFSCNSAEVPAPVK----- 198
Db 166 SSTPVVNPVKTYIGGAKVGLGTAPGPDVPSIINPTSATFDKNVTK-QADVKTMTNLGN 224
```





[4]  
RN CARBOHYDRATE-LINKAGE SITES, AND SEQUENCE OF 45-125.  
RC TISSUE=Submaxillary gland;  
RA MEDLIN=97248516; PubMed=9092502;  
RX Gerken T.A., Owens C.L., Pasumarthy M.;  
RT "Determination of the site-specific O-glycosylation pattern of the  
FT porcine submaxillary mucin tandem repeat glycopeptide. Model proposed  
FT for the polypeptide:galnac transferase peptide binding site.";  
RL J. Biol. Chem. 272:9709-9719(1997).  
CC -!- FUNCTION: APOMUCIN IS PART OF MUCIN, THE MAJOR GLYCOPROTEIN  
CC SYNTHESIZED AND SECRETED BY MUCOUS CELLS OF THE SUBMAXILLARY  
CC GLAND. ITS HIGHLY VISCOUS AQUEOUS SOLUTIONS SERVE TO LUBRICATE  
CC THE ORAL CAVITY AND TO PROTECT IT FROM THE EXTERNAL  
CC ENVIRONMENT.  
CC -!- SUBUNIT: INTERMOLECULAR DISULFIDE BONDS COULD HELP MAINTAIN A  
CC MULTIMERIC MUCIN STRUCTURE.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- TISSUE SPECIFICITY: SUBMAXILLARY MUCOSAE.  
CC -!- DOMAIN: CONTAINS TANDEMLY REPEATED, IDENTICAL SEQUENCES OF 81  
CC RESIDUES.  
CC -!- PTM: EXTENSIVELY O-LINKED GLYCOSYLATED ON SER AND THR RESIDUES OF  
CC THE REPEAT UNITS, HIGHEST GLYCOSYLATION APPEARS TO OCCUR ON SER  
CC RESIDUES WHICH HAVE GLY AT POSITIONS AT +2 OR -2 FROM THE  
CC GLYCOSYLATION SITE OR, WHERE GLY IS THE PENULTIMATE RESIDUE. THE  
CC PRESENCE OF PROLINE (USUALLY AT POSITION +3 OR -3) APPEARS TO ALSO  
CC ENHANCE GLYCOSYLATION.  
CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M61883; AAA30998.1; -;  
DR EMBL; M21174; AAA30990.1; -;  
DR PIR; A40009; A40009;  
DR InterPro; IPR000359; -;  
DR InterPro; IPR001007; -;  
DR Pfam; PF00007; Cys\_knot; 1.  
DR PROSITE; PS01185; CTCK\_1; 1.  
DR PROSITE; PS01225; CTCK\_2; 1.  
DR PROSITE; PS01208; VWFC; 1.  
KW Glycoprotein; Repeat.  
FT NON\_TER 1  
FT DOMAIN <1 368 81 AA TANDEM REPEATS.  
FT REPEAT <1 44 1.  
FT REPEAT 45 125 2.  
FT REPEAT 126 206 3.  
FT REPEAT 207 287 4.  
FT REPEAT 288 368 5.  
FT REPEAT 369 391 6 (INCOMPLETE).  
FT DOMAIN 929 995 VWFC.  
FT DOMAIN 1062 1146 CTCK.  
FT DISULFID 1062 1109 BY SIMILARITY.  
FT DISULFID 1076 1123 BY SIMILARITY.  
FT DISULFID 1085 1139 BY SIMILARITY.  
FT DISULFID 1089 1141 BY SIMILARITY.  
FT DISULFID ? 1145 BY SIMILARITY.  
FT CARBOHYD 46 46 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 50 50 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 51 51 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 57 57 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 58 58 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 61 61 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 66 66 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 67 67 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 73 73 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 74 74 O-LINKED (GALNAC. .) (MUCIN TYPE).  
FT CARBOHYD 76 76 O-LINKED (GALNAC. .) (MUCIN TYPE).  
\*

FT	CARBOHYD	77	77	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	81	81	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	83	83	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	87	87	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	91	91	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	93	93	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	94	94	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	96	96	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	98	98	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	101	101	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	103	103	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	104	104	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	106	106	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	107	107	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	108	108	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	110	110	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	111	111	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	117	117	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	123	123	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	124	124	O-LINKED (GALNAC. .)	(MUCIN TYPE).
FT	CARBOHYD	418	418	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	917	917	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	985	985	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1002	1002	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1068	1068	N-LINKED (GLCNAC. .)	(POTENTIAL).
SQ	SEQUENCE	1150 AA;	109615 MW;	3CB68B5D29DD7F5A	CRC64;

Query Match 3.6%; Score 325.5; DB 1; Length 1150;  
Best Local Similarity 20.2%; Pred. No. 6.5e-08;  
Matches 267; Conservative 148; Mismatches 451; Indels 453; Gaps 57;

Qy	70	DTTPTNSNSSSSSETASVEDSDSTTTTPDKGGAFYNAHSGLVFWFRSGTEGSLT	129
Db	1	ETARP-----SVAGSGTTGVSGAGSGTSGSGSGTG-----ATGASIGQPE	41
Qy	130	LSEIKMTGEGGAIFSGQELLFTDLTSLTIQNNLSLGGAIFG-GSTISLSGIKATFSC	188
Db	42	TSRISVAGSSGA-----PAVSSGASQAAGTSGAGPGTTASSVGVTEA---84	
Qy	189	NSAEVPAPVKKPTEPKAQTASSETSGSSSSGNDSSVSSPSSSRAEPAANLOSHTCATAT	248
Db	85	-----RPSVAGSGTTGVSGAGSGTSGSS-GSPG-----AT114	
Qy	249	PAAQTDTESTPSHPKPGSGGAIYAKDLTIADSQEVLFSINKATKGGAIKAEKDVSPEN	308
Db	115	GASIGQPETSRIS-VAGSSGA-----134	
Qy	309	ITSLKVQTNGBAEKGGAIYAKGDLTSQSSKQSLFNSNYSKOGGALYVEGGINFQDLBEI	368
Db	135	-----PAVSSGASQAAGT-----SGAGPGT-----154	
Qy	369	RIKYNKAGTFETKKTLPSPKQAQAGNADAWASSPOSGSGGATTVSDSGSSGSDST	428
Db	155	--TASSVGVTEAR---PSVAGSGTTGV-----SGA---SGSTSSSGSPGAT	195
Qy	429	SETV--PVTAKGGGLYTKNLSITNITGIIETANNKATDVGGGAYVKGTLTCENSHRLQF	486
Db	196	GASIGQPETSR-----ISVAGSSG-----APAVSSGA-----222	
Qy	487	LKNSSDKGGGIYGEDNITLSNLTGTLFQENTAKEGGGLFIKGTDKALMTGLDLSFCL	546
Db	223	-----SQAAGTSGAGPGTTASSVGVTEARPS-----VAGSGTGVSGAS---263	
Qy	547	INNTSEKHGGGAFVTKETISQYTSVDETIPGTPVHGETVITGNKS----TGNGGGVCT	602
Db	264	-GSTSSSGSPGATGASIGQPETSRIS---SVAGSSGAPAVSSGASQAAGTSGAGPGTTA	318
Qy	603	KRLALSNLQSISSISGNSAAENGGAHTCFDPSPTADTAQPPAAASAAATSTPKSPAVSTAL	662
Db	319	SSVGVTEARPSVAGSGTTGVSGA-----SGSTSSSGSPGATGASIGQPETSRISVAG	373



```
QY 663 STPSSSTVSSLTLLAASSQAQSPATSNKETQDPNADTDLLIDVVDTTISKNTAKKGGIY 722
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
374 SSGAP-----AVSGASQAAGTSEA-----TTSIEGAGTSGVGFK 408
QY 723 AKKAKMSRIDQLNISENSATEIG---GGICCKESLELDALVLSVTENLVKEG- 773
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
409 TEATTFP-----GENETTRVGIATGTGTVISRKLEPGSYNTEATTS--IGRSGTTHT 459
QY 774 ---GGLHAKTVNISNLKSGFSPNNKANSSSTGCVATTASAPAAAAAASLQAAAAAAPSSPA 830
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
460 DLPGG-----TTIVLPGFHS---SQSSKPGSSVT-----TPGSPE 492
QY 831 TPTYSG-----VVGAIYCEKYTF-----SQSGTCQFSGNQADIN 866
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
493 SGSETGTGFEFTTVISGSSHTTEATTFIGGSGPGTGRPGTTGELSGTTIASGNATTEA 552
QY 867 NPS-OSSLNVOGGAIFYAKTSLSIGSSDAGTSYIFSGNSVSTCKSQTTGOIAGGAIYSPIV 925
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
553 TTSSTRIGPQTGAQTTVPGSQVSGSETGTSEAVNPATIASG-SSTGTTSAS----- 605
QY 926 TLNCPATFSNNTASATPKTSEDSSGNSIKDITIGGAIAAGTAITLSGVSRFSNGNTADLG 985
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
606 -----DSQVTGSRGTGT-----GVVLGTTV-----APGSSSTG 633
QY 986 AAGTLANANTPSATSGSONSTKITEITLNGSFIFERNOANKRGAIYSPSVSIK-GNNIT 1044
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
634 ATTGVLINEGTRSTSLGT-TRVASGTTYESGT-----SNSVPGGSGTSGGINTGGSST 687
QY 1045 FNQNTSTHDGSAIYTKDATIESLGSVLFTGNVNTATQASSATSGQNTNTANYGRAIFGD 1104
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
688 --QVTGIOTGT-----TAVGFGSTLLPGSSNT-----GATTSPSERTS----- 723
QY 1105 PGTTQSSQTDAILTLIASSGNITFNNNSLQNNQGDTPAKFCFSIAGYVKLSLQAARKGTI 1164
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
724 PGSKGTITRVVSGTIVASGSSNTGATTSL--GRGETTQGGIKIVITGVTVGTVPAG--- 778
QY 1165 SFFDCVHTSTKKTGTQNVYETLIDINKENSNPYGTIVFSSSELHENKSYIPONAILHNG 1224
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
779 SP-----NTRATPT-----EVRAATGAGTAVGATSRSTGISTGPNST--PG 819
QY 1225 TLVLAKETELHVVSFEQKEGSKLIMEPGAVLSNQNIANGALAINGLTIDLSSMGTPQAGE 1284
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
820 T-----TEIGSGTSSPGVKYTEATTFKG---VGTTEAGISSGNSPGSGG 861
QY 1285 IFSPPF--LRIVATTS-----SASGGSGVSSSIPTNPKRISAAPVPSGSAATPTMSENK 1336
Db : : | | | | | | : : : : : : : : : : : : : : : : : : : :
862 VTSSQEGTSREASETTAPRISATGTSVSKETITASPKVSSPETTAG--ATEDQENENK 918
```

Search completed: October 25, 2001, 08:43:00  
Job time: 194 sec

**THIS PAGE BLANK (USPTO)**



PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -  
XX  
PS Claim 1; Pages 177-179; 256pp; English.  
XX  
CC The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamidiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a nucleic acid sequence  
CC isolated in the present invention.  
XX  
SQ Sequence 5265 BP; 1625 A; 1207 C; 1023 G; 1410 T; 0 other;

Query Match 100.0%; Score 5265; DB 21; Length 5265;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gcaatcatgaaatgctcagctactcgggtgttctgtgtctgtctccctcagtttca 60  
Db 1 gcaatcatgaaatgctcagctactcgggtgttctgtgtctgtctccctcagtttca 60  
Qy 61 ggggtttctccagaaacctaaagaatttaattctctcgcgtagaacctttccctct 120  
Db 61 ggggtttctccagaaacctaaagaatttaattctctcgcgtagaacctttccctct 120  
Qy 121 accacttttactgaacaattggagaagctgggcagaaatatatcgtctctggtaacgca 180  
Db 121 accacttttactgaacaattggagaagctgggcagaaatatatcgtctctggtaacgca 180  
Qy 181 tctttcacaaaatttacaacaattctactaccgatacaacaactcccacgaactcaaac 240  
Db 181 tctttcacaaaatttacaacaattctactaccgatacaacaactcccacgaactcaaac 240  
Qy 241 tctctagctctagcggagaactctccgtttctcgtaggatagtgactctacaacaag 300  
Db 241 tctctagctctagcggagaactctccgtttctcgtaggatagtgactctacaacaag 300  
Qy 301 actcctgatcctaaagggtggcgcccttttataacgcgcactccggagtttctccttt 360  
Db 301 actcctgatcctaaagggtggcgcccttttataacgcgcactccggagtttctccttt 360  
Qy 361 atgacacgatcaggaaagaaggttccttaactctgtctgagataaaatgactggtgaa 420  
Db 361 atgacacgatcaggaaagaaggttccttaactctgtctgagataaaatgactggtgaa 420  
Qy 421 ggcggtgctatctctcacaagagagctgctatttacagatctgacaagttcaaccatc 480  
Db 421 ggcggtgctatctctcacaagagagctgctatttacagatctgacaagttcaaccatc 480  
Qy 481 caaaaataactatccagctatccggagagcgatttttgaggagatctacaatctcccta 540  
Db 481 caaaaataactatccagctatccggagagcgatttttgaggagatctacaatctcccta 540  
Qy 541 tcaggattactaaagcagctttctcgtcaactctgcagaagttccctgctcctttaag 600  
Db 541 tcaggattactaaagcagctttctcgtcaactctgcagaagttccctgctcctttaag 600  
Qy 601 aaacctacagaacctaaagctcaaacagcgaacgctgggttctagttctctagc 660  
Db 601 aaacctacagaacctaaagctcaaacagcgaacgctgggttctagttctctagc 660  
Qy 661 ggaatatgattcgtgtcttccccccagttccacgtagagctgaacccgcagcagtaattt 720

Db 661 ggaatatgattcgtgtcttccccccagttccacgtagagctgaacccgcagcagtaattt 720  
Qy 721 caaagtcactttatttggctacagctactcgtcgtcctaaacccgatacacgaaacatca 780  
Db 721 caaagtcactttatttggctacagctactcgtcgtcctaaacccgatacacgaaacatca 780  
Qy 781 actcctctcctaagccagagatctggggagctactctgctaaagcgacacttactatc 840  
Db 781 actcctctcctaagccagagatctggggagctactctgctaaagcgacacttactatc 840  
Qy 841 gcagactctcaaggggtactattctcctaataaagctactaaagctggaggagcgatc 900  
Db 841 gcagactctcaaggggtactattctcctaataaagctactaaagctggaggagcgatc 900  
Qy 901 tttgctgaaagatgtttcttcgcagaattattacattataaaagtaacaaactaacgt 960  
Db 901 tttgctgaaagatgtttcttcgcagaattattacattataaaagtaacaaactaacgt 960  
Qy 961 gctgaagaaaaggagcgatctatctgctaaagtgacctctcaattcaattcttctaaa 1020  
Db 961 gctgaagaaaaggagcgatctatctgctaaagtgacctctcaattcaattcttctaaa 1020  
Qy 1021 cagagctcttttaattctaaactacagtaacaaagtggggggctctatatgttgaaagca 1080  
Db 1021 cagagctcttttaattctaaactacagtaacaaagtggggggctctatatgttgaaagca 1080  
Qy 1081 ggtataaaacttccaagatctctgaaagaatttcgcattaaagtaacaaactggaacgttc 1140  
Db 1081 ggtataaaacttccaagatctctgaaagaatttcgcattaaagtaacaaactggaacgttc 1140  
Qy 1141 gaaacaaaaaaactcaactttacctctttaaagctcaagcatctgcaggaaatgcagat 1200  
Db 1141 gaaacaaaaaaactcaactttacctctttaaagctcaagcatctgcaggaaatgcagat 1200  
Qy 1201 gcttgggctctctcctcctcaatctggtgtctgagcaactcacagttcccgactcagga 1260  
Db 1201 gcttgggctctctcctcctcaatctggtgtctgagcaactcacagttcccgactcagga 1260  
Qy 1261 gactctagctctgctcagactcggatcacctcagaaacagttccagtcacagctaaagcc 1320  
Db 1261 gactctagctctgctcagactcggatcacctcagaaacagttccagtcacagctaaagcc 1320  
Qy 1321 ggtgggcttttactgataaagaattcttcgattactaacatcacaggaattatcgaatt 1380  
Db 1321 ggtgggcttttactgataaagaattcttcgattactaacatcacaggaattatcgaatt 1380  
Qy 1381 gcaataacaaagcagacagatgttgaggtgtgctacgtataaaggaacccctacttgt 1440  
Db 1381 gcaataacaaagcagacagatgttgaggtgtgctacgtataaaggaacccctacttgt 1440  
Qy 1441 gaaaactctaccgctctacaatttttgaanaactctccgataaaacaggttgagggaatc 1500  
Db 1441 gaaaactctaccgctctacaatttttgaanaactctccgataaaacaggttgagggaatc 1500  
Qy 1501 tacggagaagaacaactcacccctatctaatgtgacaggaagactctatcccaagagaat 1560  
Db 1501 tacggagaagaacaactcacccctatctaatgtgacaggaagactctatcccaagagaat 1560  
Qy 1561 actgccaagaagagcggttgactcttcaataaaggttacagataaaagctcttacaatg 1620  
Db 1561 actgccaagaagagcggttgactcttcaataaaggttacagataaaagctcttacaatg 1620  
Qy 1621 acagagctggatagttctctttaaataacacatcacgaaaaacatgggttggaacc 1680  
Db 1621 acagagctggatagttctctttaaataacacatcacgaaaaacatgggttggaacc 1680  
Qy 1681 ttgtttacaaaagaactctctcagacttacacctctgatgtggaacaaattccaggaatc 1740  
Db 1681 ttgtttacaaaagaactctctcagacttacacctctgatgtggaacaaattccaggaatc 1740  
Qy 1741 acgctctgtacatggtgaaacagctctactctggcaataaactcacaggaggttaattggtgga 1800  
Db 1741 acgctctgtacatggtgaaacagctctactctggcaataaactcacaggaggttaattggtgga 1800

Db 1741 acgcgttacatggtgaacagtcattacttgcggcaataaaatctacaggaggaataatggtgga 1800  
Qy 1801 ggcgtgtgtacaaaacgtcttgcttatctaaaccttcaaacgttcttctatatctcgggaat 1860  
Db 1801 ggcgtgtgtacaaaacgtcttgcttatctaaaccttcaaacgttcttctatatctcgggaat 1860  
Qy 1861 tctgcgcagaaaaatggtggtgagcccaacacatgccagatagcttcccaacggcgat 1920  
Db 1861 tctgcgcagaaaaatggtggtgagcccaacacatgccagatagcttcccaacggcgat 1920  
Qy 1921 actgcagaacagccgcagcagcttctgcgcgacgtctactctactcccaaatctgcccggtc 1980  
Db 1921 actgcagaacagccgcagcagcttctgcgcgacgtctactctactcccaaatctgcccggtc 1980  
Qy 1981 tcaactgcttaagcacaccttcatcttcaacgtctcttcaacttaaccttactagagcc 2040  
Db 1981 tcaactgcttaagcacaccttcatcttcaacgtctcttcaacttaaccttactagagcc 2040  
Qy 2041 tcttcaacgctctctctgcgaacctctaaataaggaacctcaagatcctaatgctgataca 2100  
Db 2041 tcttcaacgctctctctgcgaacctctaaataaggaacctcaagatcctaatgctgataca 2100  
Qy 2101 gacttattgaatcattatgattgattagatcactatcagcaaaaacactgctaagaagcc 2160  
Db 2101 gacttattgaatcattatgattgattagatcactatcagcaaaaacactgctaagaagcc 2160  
Qy 2161 ggtggaatctatgtctaaaaagccaagatgtcccgcatagcacaaactgataatctctgag 2220  
Db 2161 ggtggaatctatgtctaaaaagccaagatgtcccgcatagcacaaactgataatctctgag 2220  
Qy 2221 aactcgcgtacagagataggtgaggttatctgctgtaagaaactctttagaactagatgct 2280  
Db 2221 aactcgcgtacagagataggtgaggttatctgctgtaagaaactctttagaactagatgct 2280  
Qy 2281 ctagtctcttctcttaacagagaacctgttggaaagaaggtggaggtctacatgct 2340  
Db 2281 ctagtctcttctcttaacagagaacctgttggaaagaaggtggaggtctacatgct 2340  
Qy 2341 aaaaactgaaatttctaatctgaaatcaggcttctcttctcgacaacaaagcaaac 2400  
Db 2341 aaaaactgaaatttctaatctgaaatcaggcttctcttctcgacaacaaagcaaac 2400  
Qy 2401 tctctatccaagaggtgcgaacaaacagcttcagcaacctgctgcagctgctgtcttcccta 2460  
Db 2401 tctctatccaagaggtgcgaacaaacagcttcagcaacctgctgcagctgctgtcttcccta 2460  
Qy 2461 caagcagccgacgacgcacacatcatctccagcaaccaacttattcaggtgtagta 2520  
Db 2461 caagcagccgacgacgcacacatcatctccagcaaccaacttattcaggtgtagta 2520  
Qy 2521 ggaggagctatctatgggaaaaaggttacattctctcaatgtagcgggacttgcagttc 2580  
Db 2521 ggaggagctatctatgggaaaaaggttacattctctcaatgtagcgggacttgcagttc 2580  
Qy 2581 tctggaaacaaagctatcgataaacaatccctcccaatcatcgtgtgaacgtacaaggagga 2640  
Db 2581 tctggaaacaaagctatcgataaacaatccctcccaatcatcgtgtgaacgtacaaggagga 2640  
Qy 2641 gccatctatgcaaaaaccttctgtctattggatcttcgatctcgatgctgaaacctctatt 2700  
Db 2641 gccatctatgcaaaaaccttctgtctattggatcttcgatctcgatgctgaaacctctatt 2700  
Qy 2701 tctcggggaacagtgcttccactgggaaatctcaaacacagggcaaatagcgggagga 2760  
Db 2701 tctcggggaacagtgcttccactgggaaatctcaaacacagggcaaatagcgggagga 2760  
Qy 2761 ggaatctatccctactgttcaattgaattgtccctgcgacatctcttaacaatacagcc 2820  
Db 2761 ggaatctatccctactgttcaattgaattgtccctgcgacatctcttaacaatacagcc 2820  
Qy 2821 tctatagctacaccggaagcttctctgaagatggatctctcaggaataatcttattaaagat 2880  
Db 2821 tctatagctacaccggaagcttctctgaagatggatctctcaggaataatcttattaaagat 2880

Qy 2881 accattggaggagccattgcaggagcagccattaccctatcttgagctctctcgatttca 2940  
Db 2881 accattggaggagccattgcaggagcagccattaccctatcttgagctctctcgatttca 2940  
Qy 2941 gggaaatacggctgatttaggagctgcaataggaaactctagctaatgcaaatacaccagt 3000  
Db 2941 gggaaatacggctgatttaggagctgcaataggaaactctagctaatgcaaatacaccagt 3000  
Qy 3001 gcaactagcggatctcaaatagcattacagaaaaatactttagaaaaacggtctcttt 3060  
Db 3001 gcaactagcggatctcaaatagcattacagaaaaatactttagaaaaacggtctcttt 3060  
Qy 3061 attttgaaagaaacaaagctaataacgtggagcgatttactctcttagcgttttcatt 3120  
Db 3061 attttgaaagaaacaaagctaataacgtggagcgatttactctcttagcgttttcatt 3120  
Qy 3121 aaagggaataataattaccttcaatcaaaaatacatccaactcatgatggaagcgtctatctac 3180  
Db 3121 aaagggaataataattaccttcaatcaaaaatacatccaactcatgatggaagcgtctatctac 3180  
Qy 3181 tttaaaaaagatgctacgattgagttcttttaggattctgttttttttacagaaaaatacgtt 3240  
Db 3181 tttaaaaaagatgctacgattgagttcttttaggattctgttttttttacagaaaaatacgtt 3240  
Qy 3241 acagctacacaagctagtctgcaacatctggcaaaaatacaaaatactgccaactatggg 3300  
Db 3241 acagctacacaagctagtctgcaacatctggcaaaaatacaaaatactgccaactatggg 3300  
Qy 3301 gcaagcaatctttggaatccaggaaacacactcaatctgctcaaacagatgcaattttaacc 3360  
Db 3301 gcaagcaatctttggaatccaggaaacacactcaatctgctcaaacagatgcaattttaacc 3360  
Qy 3361 ctctctgcttctctggaaacattacttttagcaaacagttttacagaataaccaaggt 3420  
Db 3361 ctctctgcttctctggaaacattacttttagcaaacagttttacagaataaccaaggt 3420  
Qy 3421 gatactccgcgtcagcaagtttttagtattgaggaacgtcagctcaaaactctctcacaagcc 3480  
Db 3421 gatactccgcgtcagcaagtttttagtattgaggaacgtcagctcaaaactctctcacaagcc 3480  
Qy 3481 gctaagggaagactatttagctttttogattgtgtgcacacctctaccaaaaaacaggt 3540  
Db 3481 gctaagggaagactatttagctttttttogattgtgtgcacacctctaccaaaaaacaggt 3540  
Qy 3541 tcaacacaaaacgtttatgaaacttttagattataaataagaaacagacagatcctat 3600  
Db 3541 tcaacacaaaacgtttatgaaacttttagattataaataagaaacagacagatcctat 3600  
Qy 3601 acaggaaactattgttctctcttgaaattacatgaaacaaaatcttacctccacacaaat 3660  
Db 3601 acaggaaactattgttctctcttgaaattacatgaaacaaaatcttacctccacacaaat 3660  
Qy 3661 gcaatcttcaaacggaaacttttagttctttaaagagaaaaacagaactccacgtagtctct 3720  
Db 3661 gcaatcttcaaacggaaacttttagttctttaaagagaaaaacagaactccacgtagtctct 3720  
Qy 3721 tttagcagaagaagaaggtctaaatttaattatgaaacccggagctgtgttatctaaccaa 3780  
Db 3721 tttagcagaagaagaaggtctaaatttaattatgaaacccggagctgtgttatctaaccaa 3780  
Qy 3781 aacatgactaacgcgactctagctatacaatgggttaacgattgatctttccagtatgggg 3840  
Db 3781 aacatgactaacgcgactctagctatacaatgggttaacgattgatctttccagtatgggg 3840  
Qy 3841 actctcaagcaggggaaaaatcttctctctccagaaattacgtatcgttgcacgaactct 3900  
Db 3841 actctcaagcaggggaaaaatcttctctctccagaaattacgtatcgttgcacgaactct 3900  
Qy 3901 agtgcacacaggaagcgggtcagcagtagtataccaacaataccttaaaagattctct 3960  
Db 3901 agtgcacacaggaagcgggtcagcagtagtataccaacaataccttaaaagattctct 3960

QY 3961 gcagcagtccttcagggttcgcccgaactactcgaactatgagcgagaaacaaagtcttc 4020  
Db 3961 gcagcagtccttcagggttcgcccgaactactcgaactatgagcgagaaacaaagtcttc 4020  
QY 4021 ctacacagagacatttactttaaataagatcctaataagaaacttttaccacaaacctatgttta 4080  
Db 4021 ctacacagagacatttactttaaataagatcctaataagaaacttttaccacaaacctatgttta 4080  
QY 4081 ggaagcgatcagatgtaccactaattaaagcttcgactcaacacacaaagtgagctccaaagtc 4140  
Db 4081 ggaagcgatcagatgtaccactaattaaagcttcgactcaacacacaaagtgagctccaaagtc 4140  
QY 4141 tatgattttaaacttttatctggggtatctttccctcagaagggatcatatgggaacctggaca 4200  
Db 4141 tatgattttaaacttttatctggggtatctttccctcagaagggatcatatgggaacctggaca 4200  
QY 4201 ttagattttaaacttccacaaacagggaaacttcaagccagatggacattcgatcacctatcgt 4260  
Db 4201 ttagattttaaacttccacaaacagggaaacttcaagccagatggacattcgatcacctatcgt 4260  
QY 4261 cgtcgggtatcatcatccttagggatcaatttttatgcaacctctatcttaggtcccca 4320  
Db 4261 cgtcgggtatcatcatccttagggatcaatttttatgcaacctctatcttaggtcccca 4320  
QY 4321 aactcaatgattgtgtgaagcaagggtcttacaacaacatgttgaataatgcccgcttc 4380  
Db 4321 aactcaatgattgtgtgaagcaagggtcttacaacaacatgttgaataatgcccgcttc 4380  
QY 4381 gatgatctcgttacaataacttctgggtttcagagtaggaactttcttagctcaacaa 4440  
Db 4381 gatgatctcgttacaataacttctgggtttcagagtaggaactttcttagctcaacaa 4440  
QY 4441 ggaactctcttccgaagaattcagttactacacgccgcaaaccttcagttgccatcat 4500  
Db 4441 ggaactctcttccgaagaattcagttactacacgccgcaaaccttcagttgccatcat 4500  
QY 4501 gccaaacctagacaagattttatctcagagctgcatattagtaagatagtggggaaaacc 4560  
Db 4501 gccaaacctagacaagattttatctcagagctgcatattagtaagatagtggggaaaacc 4560  
QY 4561 aaagccatacaaaaatgcatattacttccataaagggtctgagtaactcttaccacagct 4620  
Db 4561 aaagccatacaaaaatgcatattacttccataaagggtctgagtaactcttaccacagct 4620  
QY 4621 tctgtctatggaggtcaaatctctgtatttcttgcataaagcaacatggttgggcaactt 4680  
Db 4621 tctgtctatggaggtcaaatctctgtatttcttgcataaagcaacatggttgggcaactt 4680  
QY 4681 cctttcctaatacagaagggtcgttctcctatggacatattaaacatgatacaacaacactt 4740  
Db 4681 cctttcctaatacagaagggtcgttctcctatggacatattaaacatgatacaacaacactt 4740  
QY 4741 taacctctatccatgaaagaataaagagagattgggaatttaggattggttagcgat 4800  
Db 4741 taacctctatccatgaaagaataaagagagattgggaatttaggattggttagcgat 4800  
QY 4801 ctctgtatctctatgattcttaagaaccttcaagattcttctaaacggatcactgtc 4860  
Db 4801 ctctgtatctctatgattcttaagaaccttcaagattcttctaaacggatcactgtc 4860  
QY 4861 tatggggaactcagatattccagcattcgcagaaacagttccacagaaatcgattacatg 4920  
Db 4861 tatggggaactcagatattccagcattcgcagaaacagttccacagaaatcgattacatg 4920  
QY 4921 ccaagacacttcagatgattgtcttacagaaatctgcgtcttcctgtggggtcgtcgtc 4980  
Db 4921 ccaagacacttcagatgattgtcttacagaaatctgcgtcttcctgtggggtcgtcgtc 4980  
QY 4981 gaaggagctcatcatgaactgttaattcttcttgataataaagcttgcattagcctcacatg 5040  
Db 4981 gaaggagctcatcatgaactgttaattcttcttgataataaagcttgcattagcctcacatg 5040  
QY 5041 ccttctatctacagaaataactcgtctctgttaataatcgggtattgtcttcgaatgaagct 5100

Db 5041 ccttctatctacagaaataactcgtctctgttaataatcgggtattgtcttcgaatgaagct 5100  
QY 5101 ggtcaagttatctcggagtagtccaaactagaacacctctgtagacagaataacagtactcaa 5160  
Db 5101 ggtcaagttatctcggagtagtccaaactagaacacctctgtagacagaataacagtactcaa 5160  
QY 5161 ctatatcttggctccctcttgactctctacgaaactatactatcgtatgtaggcatgtat 5220  
Db 5161 ctatatcttggctccctcttgactctctacgaaactatactatcgtatgtaggcatgtat 5220  
QY 5221 acgctatcgcaaatgactagctgggtgctgcgcgatgtcttcttaa 5265  
Db 5221 acgctatcgcaaatgactagctgggtgctgcgcgatgtcttcttaa 5265

RESULT 2

AAZ01425/c  
ID AAZ01425 standard; DNA; 1038602 BP.

XX AC AAZ01425;

XX XX 07-OCT-1999 (first entry)

XX Complete genome sequence of Chlamydia trachomatis.

XX Vaccine: eye disease; conventional trachoma; nonendemic trachoma;  
XX paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;  
XX nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
XX bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.

OS Chlamydia trachomatis.

XX PN WO9928475-A2.

XX XX 10-JUN-1999.

XX XX 27-NOV-1998; 98WO-IB01939.

XX PR 04-NOV-1998; 98US-0107077

XX PR 28-NOV-1997; 97US-0095041

XX PR 17-DEC-1997; 97US-0096094

XX PA (GEST ) GENSET.

XX XX Griffais R;

XX DR WPI; 1999-371125/31.

XX PT Genome sequence of Chlamydia trachomatis

XX PS Claim 1; Page 373-656; 1755pp; English.

XX The present sequence represents the complete genome of Chlamydia  
CC trachomatis. Open reading frames (ORFs) of the genome encode  
CC polypeptides AA376754-Y37949. The polypeptides can be used as vaccines  
CC against Chlamydia trachomatis. Antisense and ribozyme sequences can also  
CC be used to control growth of the microorganism. Chlamydia trachomatis is  
CC responsible for a large number of diseases, e.g. eye diseases such as  
CC conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion  
CC conjunctivitis; genital diseases such as nongonococcal urethritis,  
CC epididymitis, cervicitis, salpingitis, perihhepatitis, bartholinitis;  
CC pneumopathy in breast feeding infants; and venereal  
CC lymphogranulomatosis. The polypeptides of the invention may be of use in  
CC treating these diseases.

XX SQ Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;

Query Match

Best Local Similarity 95.0%; Score 4999.2; DB 20; Length 1038602;

Matches 5107; Conservative 0; Mismatches 74; Indels 17; Gaps 5;



Dbb 1036445 CTACAGAGATAGGTGGAGGTATCTGCTGTAAAAAATCTTTAAAAACTAGATGCTCTAGTCT 1036386  
QY 2288 ccttatctgtaacagagaaaccttggttggaagaaggtggaggttacatcttaacataaactg 2347  
Dbb 1036385 CCTTATCTGTAACAGAGAACTTGTGGGAAAGAGGTGGAGGCTTACATGCTAAAACTG 1036326  
QY 2348 taaatatttctaactgaaatcagagcttctcttctcgaacaaagaacaaactcctcat 2407  
Dbb 1036325 TAAATATTTCTAATCTGAATCAGGCTTCTTTCTCGAACAACAAGAACTCCCTCAT 1036266  
QY 2408 ccacaggagtcgcacaacacagcttcagcaactgctgcagctgcgtctccctacaaagcag 2467  
Dbb 1036265 CCACAGAGGTGCGAACACAGCTTTCACACAGCTGCTGAGCTGCTGCTCCCTCAACAGAG 1036206  
QY 2468 ccgcagcagccgcacccatcatctccagcaacacacaaacttattcaggtgtagtagagagag 2527  
Dbb 1036205 CCGCAGCAGCGCCACCATCAATCTCCAGCAACACCAACTTATTCAGGTGTAGTAGGAGGAG 1036146  
QY 2528 ctatctatggagaaaggttacattctctcaatgttagcgggacttgcagttctctctggga 2587  
Dbb 1036145 CTATCTATGGAGAACAGTTTACATTCTCTCAATGTACGGGACTTGTCAAGTTCTCTCGGA 1036086  
QY 2588 accaagctatcgataaacaatccctcccaatcatogttaaagtcacaagagagccatct 2647  
Dbb 1036085 ACCAAGCTATCGATAACAATCCCTCCCAATCATCGTTGAACGTACAAAGGAGGAGCCATCT 1036026  
QY 2648 atgccaaaacctcttctctattgagatctccgagctggaaacctcttatatttctcgg 2707  
Dbb 1036025 ATGCCAAACACCTTTGTCTATTAGATCTTCCGATGCTGGAACTCTCTATTTTCTCGG 1035966  
QY 2708 ggaacagtgctccactgggaaatctcaaaacacagggcaaatagcgggagcgatct 2767  
Dbb 1035965 GGAACAGTGTCTCCACTGGGAAATCTCAACAACACAGGCAAAATAGCGAGGACGATCT 1035906  
QY 2768 actccctactgtttacatgtaattccctgcgaaattctctaaacaatacagcctctatag 2827  
Dbb 1035905 ACTCCCTACTGTGTACATTGGAATTTGCTGCGCAATCTCTAACAATACAGCCTCTATGG 1035846  
QY 2828 ctacaccgaagacttctctgaagatgatctcaggaattctttaaagataccattg 2887  
Dbb 1035845 CTACACCGAAGACTTCTC-----ATCCTCAGGAATTTCTATTAAAGATACCAATG 1035795  
QY 2888 gaggagccattgcaggagcagccattaccctatctggtgagctctcgtatttccagggaata 2947  
Dbb 1035794 GAGGAGCCATTGACGGGACAGCCATTACCCTATCTGAGTCTCTCGATTTTCAGGGAATA 1035735  
QY 2948 cggctgatttagagcgtcaataggaaactctagctaaatgcgaatacacccagtgcaacta 3007  
Dbb 1035734 CGGCTGATTTAGGAGCTGCAATAGGAATCTTAGCTAATGCAAAATACACCCAGAGCAACTA 1035675  
QY 3008 gcggtctcaaaatagcattacagaaaaaattactttagaaaaaggttctttattttg 3067  
Dbb 1035674 GCGGATCTCAAAATAGCATTTACAGAAAAAATTACTTTTAAAAACAGTCTTTTATTTTTG 1035615  
QY 3068 aaagaaaccaaagctaaataacgtggagcgttactctctcctagcgtttccacttaagggga 3127  
Dbb 1035614 AAAGAAACCAAGCTAATAACGCTGGAGCGATTTACTCTCTCTAGCGTTTCCATTAAAGGGA 1035555  
QY 3128 ataattattaccttaacaaatacatccactcatgtaggaagcgtactctactttacaa 3187  
Dbb 1035554 ATAATATTTACCTTCAAATCAAAATACATCCACTCATGATGGAAGCGCTATCTACTTTTACAA 1035495  
QY 3188 aagatgctcacgattgagcttttagatctgttctttttacaggaataacgtttcacagcta 3247  
Dbb 1035494 AAGATGCTAAGATTGAGTCTTTTAGATCTGTTCTTTTACAGGAAATAAAGTTTACAGCTA 1035435  
QY 3248 cacaagctagttctgcaacatctggacaaaaatacaaaatcactgccaactatggtgggcagcca 3307  
Dbb 1035434 CACAAGCTAGTTCTGCAACATCTGGACAAAATACAAAATACTGCCAACTATGTGGGCAGCTA 1035375  
QY 3308 totttggagatccagaagaaacctcaatcgtctcaaaacagatgcgaattttaaccttcttg 3367  
Dbb 1035374 TCTTTGGAGATCCAGGAACCACTCAATCGCGCTCAACAGATGCCATTTTAAACCCCTTTG 1035315

QY 3368 cttcttctggaaacattacttttagcaacaacagttttacagaataaccaaggtgatactc 3427  
Dbb 1035314 CTCTCTCTGGAACATTTACTTTTAGCAACAACAGATTTTACAGATAACCAAGGTGGTACTC 1035255  
QY 3428 ccgctagcaagttttgttagtattgcaggatacgtcaaaactctctctacaagccgcftaaag 3487  
Dbb 1035254 CCGCTAGCAAGTTTGTGTAGTATTGTCAGGATACGTCAAACTCTCTTTTACAAGCCCAATAAG 1035195  
QY 3488 ggaagactattagctttttctgattgtgcacacctctacaaaaaaacaggtttcaaac 3547  
Dbb 1035194 GGAAGACTATTAGCTTTTTCGATTGTGCACACCTTACCAAAAAACAGGTTTCAACAC 1035135  
QY 3548 aaaaagctttatgaaacttttagatatataaagaagaagaaacagtaataatccatatacaggaa 3607  
Dbb 1035134 AAAACGTTTATGAACCTTTAGATATTAAATAAAGAGAGACAGTAAGCCATATACAGGAA 1035075  
QY 3608 ctattgttctctctctgaattacatgaaaaaaatcttatcccccacagaatgcaatcc 3667  
Dbb 1035074 CTATTGTGTCTCTCTGAAATTACATGAAAAAATACTTACATCCACAGAAATGCAATCC 1035015  
QY 3668 ttcacaacgcgaacttttagttttaaagaagaaaaacagaactccacagtagtctcttttggag 3727  
Dbb 1035014 TTCAACAACGGAACCTTATAGTTCTTAAAGAGAAAAACAGAACTCCACGTAGTCTCTTTTGAGC 1034955  
QY 3728 agaagaagaggttctaaaaatattatggaaacccggagctgtgttatctaaacaaaaacatac 3787  
Dbb 1034954 AGAAAGAAAGGCTCTAAATTAATTAATATGGAACCCGGAGCTGTGTTATCTAACCAAAACATAG 1034895  
QY 3788 ctaacggagctctagctatcaatgggttaacgattgatcttccagtagtgggagctctc 3847  
Dbb 1034894 CTAAACGAGCTCTAGCTATCAATGGGTTAAAGATTGATCTTTTCCAGTTATGGGACTCCTC 1034835  
QY 3848 aagcaagggaaaacttctctctccagaaattacgtatogtggcaacgacctctagtgat 3907  
Dbb 1034834 AAGCAGGGGAAATCTCTCTCCCTCCAGAAATACGATATCGTTGCCACAGCTCTAGTGCAT 1034775  
QY 3908 ccgcaggaagcgggctcagcagtagtataaccaaaaaatctctaaagaggtttctcgcagcag 3967  
Dbb 1034774 CCGGAAGAAGCGGGCTCAGCAGGAGTATACCAACAATCTTAAAGGGATTCTCGCAGCAG 1034715  
QY 3968 tgccttcagggtctgcgcgaactactccaactatgagcgagaacaagaatgttttccctaaacag 4027  
Dbb 1034714 CGCCTTCAGGTTCTGCGCGCAACTACTCCAATATGAGCGAGAGAACAAAAGTTTTCCTAACAG 1034655  
QY 4028 gagaccttactttaatacactcctaagaaacttttaccaaaacccctatggttagaaagc 4087  
Dbb 1034654 GAGACCTTACTTTAATAGATCCTAATGGAACTTTTATCAAAACCCCTATGTTAGGAAGCG 1034595  
QY 4088 atctagatgtaccactaataatgaagcttccgacttaacacaagtgacgtccaagtcttatgatt 4147  
Dbb 1034594 ATCTAGATGTACCCTAATTAAGCTTCCGACTAACACAAGTGACGTCCAAGTCTATGATT 1034535  
QY 4148 taactttatctgggagatcttttccctcagaaaggttacatgggaacctggacatttagatt 4207  
Dbb 1034534 TAACCTTATCTGGGATCTTTTCCCTCAGAAAGGTTACATGGGAACCTGGACATTAATAAT 1034475  
QY 4208 ctaatcccaacaagggaaaacttcaagccagatgacattcgatacctatcgtcgtcgtgg 4267  
Dbb 1034474 CTAATCCACAACAGGAAACTTCAAGCCAGATGGACATTGACCTATCGTCCGTGGGG 1034415  
QY 4268 tatatacacttagggataatcatttttatggaaactctatcttaggctcccaaaaactcaa 4327  
Dbb 1034414 TATACATACCTAGGGATAATCATTTTTTATGGAACTCTATCTTAGGCTCCCAAACTCAA 1034355  
QY 4328 tgaattgttgtaagcaagggcttatcaacaacatgttgtaataatgcccgcttcgatgata 4387  
Dbb 1034354 TGATTGTTGTAAGCAAGGGCTTATCAACAACATGTTGAATAATGCCGCTTCGATGATA 1034295  
QY 4388 tgcgttaacaataactctcgggttttcagagtaggaactttcttagctcaacaaggaactc 4447  
Dbb 1034294 TCGCTTACAATAACTTCTGGGTTTCAGAGTAGGAACCTTCTTAGCTCAACAAGGAACCTC 1034235



QY 4448 ctctttccgaagaattcagttactaacgacgcgcgaacttcagttgccatcgatgccaaac 4507  
DB 1034234 CTCCTTCGGAAGAAATTCAGTTACTACAGCCGCGAACTTCAGTTGCCATCGCAAAAC 1034175  
QY 4508 ctagacaagattttatctcctagagctgcatttagtaagtagtgggaaacccaagcca 4567  
DB 1034174 CTAGACAAGATTTATCTCTAGAGCTGCATTTAGTAAGATGGTGGGAAACCAAGCA 1034115  
QY 4568 tcaaaaaaatgcataattacttccataaagcgtctgagctactcttaccagcttctgtct 4627  
DB 1034114 TCAAAAAATGCATAATTACTTCCATAGGCGCTCTGAGTACTCTTACCAAGCTTCTGTCT 1034055  
QY 4628 atggaggtaaatctcctgtatttcttctgctcaataaagaacatggttgggcaactctctcc 4687  
DB 1034054 ATGGAGGTAAATTCCTGTATTCTTCTGCTCAATAAGCAACATGGTTGGGCACCTCTCTTCC 1033995  
QY 4688 taatacaggagctcgtcctctatgaacatattaaacatgatacaaacacattaccctt 4747  
DB 1033994 TAATACAAGGAGTCTGTCTCTATGGACATATTAAACATGATACAATACTTTACCCCTT 1033935  
QY 4748 ctatccatgaagaataaagagattgggaagatttaggattggttagcggtattcttcgta 4807  
DB 1033934 CTATCCATGAAGAAATAAAGAGATTGGGAAGATTTAGGATGGTTACGGGATCTTCGTA 1033875  
QY 4808 tctctatggatcttaagaacacctcttaagattcttctaaacggatcactgtctctatgggg 4867  
DB 1033874 TCTCTATGGATCTTAAAGAAGCTTCTTAAAGATTTCTTAAACGGGATCACTGCTATGGGG 1033815  
QY 4868 aactcagattccagcattcccgagaacagttcacagaatacattacgacccaagac 4927  
DB 1033814 AACTCGAGTATTCAGCAATTCGCCAGAACAGTTTACAGAAATCGATTACCATCAAGAC 1033755  
QY 4928 acttcgatgattgcttacagaaaactctgccttctctggtggtgcgctgcgaaggag 4987  
DB 1033754 ACTTCGATGATTGCTTACAGAAATCTGCTGCTTCTGCGATGCGCTGTCGAAGGAG 1033695  
QY 4988 ctatcatgaactgttaatatctctatgtataataagcttgcattagcctacatgcctcta 5047  
DB 1033694 CTATCATGAACCTGAATAATTTCTATGTATAAAGCTTGCATTTAGCCFACATGCCCTTCTA 1033635  
QY 5048 tctacagaataatcctctgtctaaatcaggtattgtcttcgaatgaagctgccaag 5107  
DB 1033634 TCTACAGAAATAATCTGCTGTAAATATCGGATTTGCTTCGAATGAAGCTGGTCAAG 1033575  
QY 5108 ttatctcggagtgccaactagaaactctgctagagcagaatacacagtactcaactatc 5167  
DB 1033574 TTATCTGGGAGTGCCAACTAGAACCTCTGCTAGAGCAGATACAGTACTCAACTATATC 1033515  
QY 5168 ttgggtccctcttgactctctacggaactatactatcgatgtaggcattgtatagcgtat 5227  
DB 1033514 TTGGTCCCTTCTGACTCTCTACGGAACATATACTATCGATGAGGCATGTATAC-CTAT 1033456  
QY 5228 cgaaatgactagctgcggtgctcgcgatgattcttaa 5265  
DB 1033455 CGAAATGACTAGCTGCGGTGCTGCGATGATCTCTCTAA 1033418

RESULT 3

AAA64755

ID AAA64755 standard; DNA; 5331 BP.

XX

AC AAA64755;

XX

DT 02-FEB-2001 (first entry)

XX C. trachomatis pmpC gene coding sequence.

XX Chlamydial infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;

KW trachoma; blindness; acute respiratory tract infection;

KW atherosclerosis; coronary heart disease; antibacterial; ss.

XX Chlamydia trachomatis.

XX WO200034483-A2.  
PN XX  
XX 15-JUN-2000.  
PD XX  
XX 08-DEC-1999; 99WO-US29012.  
PF XX  
XX 08-DEC-1998; 98US-0208277.  
PR 08-APR-1999; 99US-0288594.  
PR 01-OCT-1999; 99US-0410568.  
PR 22-OCT-1999; 99US-0426571.  
XX (CORI-) CORIXA CORP.  
PA  
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
PI WPI; 2000-431303/37.  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -  
XX  
PS Claim 1; Pages 176-177; 256pp; English.  
XX  
CC The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is a  
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a nucleic acid sequence  
CC isolated in the present invention.  
XX  
SQ Sequence 5331 BP; 1734 A; 1170 C; 1062 G; 1365 T; 0 other;

Query Match 15.8%; Score 830.6; DB 21; Length 5331;  
Best Local Similarity 59.1%; Pred. No. 5.1e-210;  
Matches 1769; Conservative 0; Mismatches 1014; Indels 210; Gaps 12;

QY 2426 cagcttcagcactgctgcagctgctgtctcctcaaacgagccgcgagcagcgacacat 2485  
DB 2396 cagactcttctcctctgagatagcgtggagactctgaagaccgactgagccagaag 2455  
QY 2486 catctccagcaacacacacttattcaggtgttagtaggagagctatctatgagaaaaag 2545  
DB 2456 ctggttctacaacagaaactcctactttaataggaggaggtgctatctatgagaaaaatg 2515  
QY 2546 ttacattctctcaatgtagcgggacttgcgtctctctgggaaccaagctatcgataaca 2605  
DB 2516 ttaagattgagaacttctctgccaaggatattttctggaaacaaagctatcgataaca 2575  
QY 2606 -----atccctcccaatcatcgttgaacgtacaagagagaccatctatgccaaaact 2659  
DB 2576 ccacagaaaggctcctcttccaaatcaacgtcctcgaggtcggttctatgctataaacat 2635  
QY 2660 ctttgtctattggatcttcgatgctggaaacctcctatatttctcgggaacagtgct 2719  
DB 2636 tgtttaatctgatagcgggagcttagcgaactgtcacccttctccgggaataactgtct 2695  
QY 2720 ccactgggaaatctcaaacacacagggcaaatagcgggagagcgatctactccctactg 2779  
DB 2696 cttct-----caatctacaacaggtcaggttgctggaggagctactactctctactg 2749  
QY 2780 ttacattgaattgctcgcgacattctctcaacaatcacagcctctatagctacacgaaga 2839  
DB 2750 taaccattgctactcctgtagttattttctaaaaactctgcaacaaactgctataaacg 2809



Db	4919	ttcaactgtttacggagaaattcggatactccagtatccgttcagaaacaattccagagaacag	4978
Qy	4913	attacgatccaagacactcgcgatgttgctttacagaaattctgcgttccttgagat	4972
Db	4979	aatacgaatcctcgttacttcgcgaactgcacataagaaacttagcaattcctatggggt	5038
Qy	4973	gcctctcgaagagactatcatgaactgtaattctttatgtataataaagcttgcattag	5032
Db	5039	tagcattcgaaggagagcctctcgtgtaacgaattcttgatgacaacagattctctgtag	5098
Qy	5033	ccacatgcctctctatctacagaaataatcctcgtgtaaaatcgcgggtattgtcttcga	5092
Db	5099	catacatgcatacaattctatcgaaattctcaaacatgcaaatcaccaagtctctcttcag	5158
Qy	5093	atgaagctggtcaagttatctcggaggtgcgaacttagaacctctcgttagagcagaataca	5152
Db	5159	gagaagcggagaaaattatttggtagtaccgcagaagaaactcagctcgcggagaataca	5218
Qy	5153	gtactcaactatatttggctcctctcggactctcgaactcgaagaaactactatcgatgtag	5212
Db	5219	gcacgcagctgtaccgggaacctttgtggactctgtatgtatcctacacagatagaacag	5278
Qy	5213	gcattgtatacgcgtatcgaaaatgactagctcgggtgcgtcgcatgatcttctaa	5265
Db	5279	acgcacaatacactgatcctatgtatgaactcgggtgcgtcgatgacattctaa	5331

## RESULT 4

AA64762	
ID	AA64762 standard; DNA; 2847 BP.
XX	
XX	AA64762;
XX	
DT	02-FEB-2001 (first entry)
XX	
DE	C. trachomatis pmpC gene carboxy terminus coding sequence.
XX	
KW	Chlamydial infection; sexually transmitted disease;
KW	pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW	trachoma; blindness; acute respiratory tract infection;
KW	atherosclerosis; coronary heart disease; antibacterial; ss.
XX	
OS	Chlamydia trachomatis.
XX	
PN	W0200034483-A2.
XX	
PD	15-JUN-2000.
XX	
XX	08-DEC-1999; 99WO-US29012.
XX	
PR	08-DEC-1998; 98US-0208277.
PR	08-APR-1999; 99US-0288594.
PR	01-OCT-1999; 99US-0410568.
PR	22-OCT-1999; 99US-0426571.
XX	
XX	(CORI-) CORIXA CORP.
PI	
PI	Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX	
XX	WPI; 2000-431303/37.
DR	
XX	
PT	Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT	comprises immunogenic portion of Chlamydia antigen, which comprises
PT	amino acid sequence encoded by polynucleotide sequence -
XX	
PS	Claim 1; Pages 203-204; 256pp; English.
XX	
CC	The present invention relates to new nucleic acid sequences and the
CC	proteins encoded by the nucleic acid sequences. The encoded proteins
CC	comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC	proteins are useful for the serodiagnosis and treatment of Chlamydia
CC	infection. Chlamydiae are intracellular bacterial pathogen that are



XX Chlamydia infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial; ss.  
XX Chlamydia trachomatis.  
XX WO200034483-A2.  
XX 15-JUN-2000.  
XX 08-DEC-1999; 99WO-US29012.  
XX 08-DEC-1998; 98US-0208277.  
XX 08-APR-1999; 99US-0288594.  
XX 01-OCT-1999; 99US-0410568.  
XX 22-OCT-1999; 99US-0426571.  
XX (CORI-) CORIXA CORP.  
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
XX WPI; 2000-431303/37.  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -  
XX Claim 1; Page 240; 256pp; English.  
XX The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a nucleic acid sequence  
CC isolated in the present invention.  
XX Sequence 598 BP; 200 A; 126 C; 123 G; 149 T; 0 other;  
XX  
XX Query Match 10.7%; Score 564; DB 21; Length 598;  
XX Best Local Similarity 98.8%; Pred. NO. 9.5e-140;  
XX Matches 589; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
QY 1209 cctctctctcctcaatcgttcttgagcaactacagctctccgactcaggagactctag 1268  
DB 1 cctctctctcctcaatcgttcttgagcaactacagctctccgactcaggagactctag 60  
QY 1269 cctcggctcagactcggatacctcagaaacagttccagtcacagctaaaggcgttggtc 1328  
DB 61 cctcggctcagactcggatacctcagaaacagttccagtcacagctaaaggcgttggtc 120  
QY 1329 ttatactgataagaattcttcgattactaatacatcacaggaattatcgaaattgcaaataa 1388  
DB 121 ttatactgataagaattcttcgattactaatacatcacaggaattatcgaaattgcaaataa 180  
QY 1389 caaagcgacagatgttgagggtggtgcttacgtacgtacgtacgtacgtacgtacgtacgt 1448  
DB 181 caaagcgacagatgttgagggtggtgcttacgtacgtacgtacgtacgtacgtacgtacgt 240  
QY 1449 tcacgcctcacatctttgaaacactctccgataaaacaggtggagggaattacagcaga 1508  
DB 241 tcacgcctcacatctttgaaacactctccgataaaacaggtggagggaattacagcaga 300

QY 1509 agacaacatcacctatcttaatttgacaggagactctattccaagaataactgcca 1568  
DB 301 agacaacatcacctatcttaatttgacaggagactctattccaagaataactgcca 360  
QY 1569 agaagagggcgttgagactcttataaaaggtacagataaagctcttacaatgacagact 1628  
DB 361 agaagagggcgttgagactcttataaaaggtacagataaagctcttacaatgacagact 420  
QY 1629 ggatagttcttcttaataataacacatcacagaaacatggtggt-ggagcctttgtta 1687  
DB 421 ggatagttcttcttaataataacacatcacagaaacatggtggtggtggtggtttgtta 480  
QY 1688 ccaagaaatctctcagacttacacctc-tgatgtggaacaaattccaggaatcacgcct 1746  
DB 481 ccaagaaatctctcagacttacacctc-tgatgtggaacaaattccaggaatcacgcct 540  
QY 1747 gacatggtgaaacagctcattactggaataataatcacaggaggttaattggtggagg 1802  
DB 541 gacatggtgaaacagctcattactggaataataatcacaggaggttaattggtggagg 596  
RESULT 6  
AAA64796/C  
ID AAA64796 standard; DNA; 511 BP.  
XX  
AC AAA64796;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE C. trachomatis 5' end CcL2gam-30 coding sequence.  
XX  
KW Chlamydial infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial; ss.  
XX Chlamydia trachomatis.  
XX OS  
XX WO200034483-A2.  
XX 15-JUN-2000.  
XX 08-DEC-1999; 99WO-US29012.  
XX 08-DEC-1998; 98US-0208277.  
XX 08-APR-1999; 99US-0288594.  
XX 01-OCT-1999; 99US-0410568.  
XX 22-OCT-1999; 99US-0426571.  
XX (CORI-) CORIXA CORP.  
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
XX WPI; 2000-431303/37.  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -  
XX Claim 1; Page 240; 256pp; English.  
XX The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a nucleic acid sequence  
CC isolated in the present invention.  
XX Sequence 598 BP; 200 A; 126 C; 123 G; 149 T; 0 other;  
XX  
XX Query Match 10.7%; Score 564; DB 21; Length 598;  
XX Best Local Similarity 98.8%; Pred. NO. 9.5e-140;  
XX Matches 589; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
QY 1209 cctctctctcctcaatcgttcttgagcaactacagctctccgactcaggagactctag 1268  
DB 1 cctctctctcctcaatcgttcttgagcaactacagctctccgactcaggagactctag 60  
QY 1269 cctcggctcagactcggatacctcagaaacagttccagtcacagctaaaggcgttggtc 1328  
DB 61 cctcggctcagactcggatacctcagaaacagttccagtcacagctaaaggcgttggtc 120  
QY 1329 ttatactgataagaattcttcgattactaatacatcacaggaattatcgaaattgcaaataa 1388  
DB 121 ttatactgataagaattcttcgattactaatacatcacaggaattatcgaaattgcaaataa 180  
QY 1389 caaagcgacagatgttgagggtggtgcttacgtacgtacgtacgtacgtacgtacgtacgt 1448  
DB 181 caaagcgacagatgttgagggtggtgcttacgtacgtacgtacgtacgtacgtacgtacgt 240  
QY 1449 tcacgcctcacatctttgaaacactctccgataaaacaggtggagggaattacagcaga 1508  
DB 241 tcacgcctcacatctttgaaacactctccgataaaacaggtggagggaattacagcaga 300

```
CC coronary heart disease. The present sequence is a nucleic acid sequence
XX isolated in the present invention.
SQ Sequence 511 BP; 132 A; 101 C; 131 G; 145 T; 2 other;

Query Match          9.0%; Score 472.2; DB 21; Length 511;
Best Local Similarity 98.8%; Pred. No. 2.1e-115;
Matches 485; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 2216 ctgagaactcgcgtacagagataggtggaggtatctgctgtaagaatactttagaactag 2275
Db 511 CTGAGAACTCGCTACAGAGATAGGTGGAGGTATCTGCTGTAAATAATCTTTAAACATAG 452

Qy 2276 atgc-tctagtcctctatctgaacagaaaccttgttggaagaaggtggagcctta 2334
Db 451 ATGCTCTAGTCTCTTAATCTGTACAGAGAACCTTGTGGGAAGAAGGTGGAGGCTTA 392

Qy 2335 catgctaaactgtaaattatttctaatactgaaatcaggtcttcttctcgaaacacaaa 2394
Db 391 CATGCTAAACTGTAAATATTTCTAATCTGAATCAGGCTTCTTCTTCGACAAACAAA 332

Qy 2395 gcaaaactcctcatcacagaggtcgcaacacagcttcagcacctgtgagctgctgt 2454
Db 331 GCAAACTCCTCATCCACAGAGTGCACAAACAGCTTCAGCACCTGCTGCAGCTGCTGCT 272

Qy 2455 tccctcaagcagcgcgcagcgcacccatcatctccagcaacacaaacttattcaggt 2514
Db 271 TCCCTTACAAAGCAGCGCAGCAGCGCACCATCATCTCCAGCAACACCAACTATTTCAGGT 212

Qy 2515 gtatgagagagctatctatggagaaaaggtttacattctctcaatgtagcgggacttgt 2574
Db 211 GTAGTAGGAGGAGCTATCTATGGAGAACAGCTTACATTCTCTCAATGTAGCGGGACTTGT 152

Qy 2575 cagttctctgggaacacagctatcgatatacaaatccctcccaatcatctgttaagatcaaa 2634
Db 151 CAGTTCTCTGGGAACCAACAGCTATCGATACCAATCCCTCCCAATCATCTGTTGAACGTACAA 92

Qy 2635 ggagagccatctatgccaaacctcttgtctattgtatcttcgatcttcgatctggaaacctcc 2694
Db 91 GGAGGAGGCATCTATGCCAAACCTCTTTGTCTATTAGATCTTCGGATGWTGGAACTCC 32

Qy 2695 tatattttctc 2705
Db 31 TATATTTTCTC 21

RESULT 7
AAA50034
ID AAA50034 standard; DNA; 1550 BP.
XX
AC AAA50034;
XX
DT 10-OCT-2000 (first entry)
XX
DE DNA encoding Chlamydia pneumoniae antigen CPN100711 RY-59.
XX
KW CPN100711 RY-59; antigen; infection; diagnosis; therapy; vaccine;
KW outer membrane protein; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 10..1419
FT /*tag= a
XX
XX WO200039158-A1.
XX
XX 06-JUL-2000.
XX
XX 23-DEC-1999; 99WO-CA01230.
XX
XX 23-DEC-1998; 98US-0113280.
```

```
PR 23-DEC-1998; 98US-0113281.
PR 23-DEC-1998; 98US-0113282.
PR 23-DEC-1998; 98US-0113283.
PR 23-DEC-1998; 98US-0113284.
PR 23-DEC-1998; 98US-0113285.
PR 23-DEC-1998; 98US-0113385.
PR 28-DEC-1998; 98US-0114050.
PR 28-DEC-1998; 98US-0114056.
PR 28-DEC-1998; 98US-0114057.
PR 28-DEC-1998; 98US-0114058.
PR 28-DEC-1998; 98US-0114059.
PR 28-DEC-1998; 98US-0114061.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
XX Murdin AD, Oomen RP, Wang J;
XX
XX WPI; 2000-452369/39.
XX P-PSDB; AAY95547.
XX
XX Novel Chlamydia polynucleotides and polypeptides useful for diagnosis,
XX prevention and treatment of Chlamydia infection in mammals -
XX
XX Claim 2(a); Page 62-65; 215pp; English.
XX
XX The present sequence is that of Chlamydia pneumoniae genomic DNA
XX including an open reading frame that codes for CPN100711 RY-59 (see
XX AAY95547), a putative outer membrane protein. It is an example of
XX C. pneumoniae polynucleotide molecules of the invention (see
XX AAA50030-42) that encode antigenic polypeptides (see AAY95543-55) useful
XX in the diagnosis, treatment and prevention of Chlamydia infection.
XX The polynucleotides can be utilised: in the recombinant production
XX of Chlamydia antigens using transformed unicellular host cells; in
XX vaccines or live vaccine vectors; in naked form or formulated with
XX a delivery vehicle for therapy and prophylaxis of Chlamydia
XX infection; and as probes or primers for diagnosis of Chlamydia
XX infection.
XX
XX Sequence 1550 BP; 467 A; 341 C; 306 G; 436 T; 0 other;

Query Match          8.3%; Score 435.8; DB 21; Length 1550;
Best Local Similarity 60.3%; Pred. No. 1.8e-105;
Matches 763; Conservative 0; Mismatches 487; Indels 15; Gaps 2;

Qy 4003 agcgagaacaaagtttcttaacagagagacaccttatttaatatagatcctaaggaaacttt 4062
Db 166 agtaagataagattgattatcacaggaactgtgactctctagatcctcaatggaactta 225

Qy 4063 taccaaaacctatgttaggaagcgatctagatgtaccactaattaaagcttccgactaac 4122
Db 226 tatcaaaattcttattctgttggaagacgcgatcatcactcttttcaatatagacaattct 285

Qy 4123 acaagtgac---gtccaagtctatgatttaactttatctggggatcttttccctcagaaa 4179
Db 286 gcaagtggggcagtttacagccacgaatgtccaccttcagggaatttaggagctaaaaaa 345

Qy 4180 gggatcatgggaaccttggaacatttagatttctaataccaaaaacagggaaacttcaagccaga 4239
Db 346 ggatatttaggaacctggaaatttggtccaattcctcggttccaaaaattattctaaaaa 405

Qy 4240 tggcatttcgatacctatcgctcgctgggtggtatcacatcactagggataatcttttttcgcg 4299
Db 406 tggacctttgacaaatacctgcgtgcctacatccctagacagacacactttctacac 465

Qy 4300 aactctatcttaggtcccaaaactcaatgatgtgttgaggaagggtctatcaacaac 4359
Db 466 aactctatttgggagcacaacactctttagtgactgtgaaaccaagggtatcttagggaac 525

Qy 4360 atgttgaataaatgccgcttcgatgatcgcttacaataaactctctgggtttcaggagta 4419
Db 526 atgttgaacaatgcaagggttgaagatcctcttccaacaacactctctgggttcggctata 585
```

```
QY 4420 ggaactttcttagctcaacaagaactctctcttccgagaagaattcagttactacagccgc 4479
DB 586 ggaactttcttaggaaagaagatctgaaattctgactcattccattcagtcagcaga 645
QY 4480 ggaacttcagttgcatcgatccaaacacacagatgacagattttatctcagagctgcatit 4539
DB 646 ggcataccgcgcgcgtggtgagccaaacccgcgaagaattttatttttagagctgcttc 705
QY 4540 agtaagatagtggtggaaaaaaccaaacgcatcaaaaaaatgacataattacttccataagggc 4599
DB 706 agtcagggttttgggtcacgcgcgagctgctgaatcaccttgacaaactaagaacataagc 765
QY 4600 tctgagtaactcttacaagaacttctgctatgaggttaaatctctgctattcttctgctcaat 4659
DB 766 tcaggctcactacacaagcatctctttatgctggaatatcttcttatcttctctcgogata 825
QY 4660 aagcaacatggttgccactctcttccataacagaagagtgctgctctatggacataatt 4719
DB 826 c-----ggctcggcctattctctatccaaaggtgctgacacattggttatg 873
QY 4720 aaacatgatacaacaacactttacccttctatccatgaagaataaagagattggaa 4779
DB 874 caacatgacacacaacacttactctctctattgaagaaaaaatgcaaaactggat 933
QY 4780 gatttagatggttagcggtattctctctctctctctctctctctctctctctctctctct 4839
DB 934 agcattgctggttattgattgctggttccagtggtggttcttaaaagaacactcaacctcac 993
QY 4840 tctctaaacgactcactgctctatggggaactcgactgattccagcattccgcagaaacag 4899
DB 994 tctacagaaggtcttacctctctacagaagctgagatcaccagaattccgcagggaaa 1053
QY 4900 ttacagaatacgattacgataccacacacttctgattgctgttctacagaaatctgtcg 4959
DB 1054 ttacagaagctagactatgactagattcttctctgctgctctcttctggaacttagca 1113
QY 4960 cttctgtggatgctgctgcaagagctatcagaaactgttaattctcttctgataat 5019
DB 1114 attcctactggattctctctgtagacggagactgctggtggtgagattctctctataat 1173
QY 5020 agcttgcttagctacactcctctctctctctctctctctctctctctctctctctctct 5079
DB 1174 aaagtacagctgctgactcctctctgctgctgctgctgctgctgctgctgctgctgct 1233
QY 5080 gtattgtcttgaatgaagctggtcaagttatctctgctgctgctgctgctgctgctgctgct 5139
DB 1234 gttctctctacaaaagaaaagggcaacgtagtcaacgttctccctacagaacacgagct 1293
QY 5140 agagcagaatacagactactcaactatattctgctccctctgactctctacaggaactat 5199
DB 1294 cgtgcagaggtgagctctcaaatcttctggaagttaactggaacactctacggcacgtat 1353
QY 5200 actatcgatgtagcatgtatcgcctatcgcaaatgactgctgctgctgctgctgctgctgct 5259
DB 1354 actatgagcttcaatgaatacttctgctgcaaatgccaagggagggatccggtgtgta 1413
QY 5260 ttcta 5264
DB 1414 ttcta 1418
```

RESULT 8

```
AAx91990
ID AAX91990 standard; DNA; 1230025 BP.
XX
AC AAX91990;
XX
DT 13-SEP-1999 (first entry)
DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
```

```
KW vaccine; neutralising epitope; ss.
XX Chlamydia pneumoniae.
XX
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-1B01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX
XX 21-NOV-1997; 97FR-0014673.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae
XX
XX Claim 1; Page 291-611; 1912pp; English.
XX
CC The present sequence represents the complete genome of Chlamydia
pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
respiratory disease such as pneumonia and bronchitis and is thought
to be a contributing factor in heart disease, sarcoidosis, sinusitis,
purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
encoded by the open reading frames of the C. pneumoniae genome (see
AAY34584-Y35879) can be used in immunogenic compositions as vaccines.
CC Vectors containing C. pneumoniae nucleotide sequences can also be
used as immunogenic compositions, especially where the vector directs
the expression of a neutralising epitope of C. pneumoniae.
XX
XX Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;
```

```
Query Match 8.3%; Score 435.8; DB 20; Length 1230025;
Best Local Similarity 60.3%; Pred. No. 7.8e-104;
Matches 763; Conservative 0; Mismatches 487; Indels 15; Gaps 2;

QY 4003 agcgagaaacaaagtttctctacagagagacaccttactttaatagatcctaatggaaacttt 4062
DB 635707 agtaagataagattgatattacaggaactgtgactctctctagatcctaatggcaactta 635766
QY 4063 taccaaaacccctatgttagaagcgcgactagatgaccactaattagcttcgcgactaac 4122
DB 635767 tatcaaaattcttatcttggtagaagaccgcgatacactcttttccaatatagacaattct 635826
QY 4123 acaagtgac---gtccaagtctatgatttaacctttatctgggagatcttttccctcagaaa 4179
DB 635827 gcaagtggggcagttacagccagaaatgcaccttcaaggggaatttagagctaaacaa 635886
QY 4180 gggatcattgggaacctggacatttagatttcaatcccaaacacaggggaaaccttcaagcaga 4239
DB 635887 ggatatttagaacctggatttggatccaaattctcgggttccaaaaattattctaaaa 635946
QY 4240 tggacattcgatacctatcgctcgtgggtatatacatccttaggagataatcttttatcg 4299
DB 635947 tggacctttgacaaatcacctgcgctgacctacatccctcagagacacacctctcatc 636006
QY 4300 aactctatcttagctcccaaaactcaatgattgttgaaagcaagggcttatcaacaac 4359
DB 636007 aactctatttgggagacacaaaactctttagtgactgtgaaacaaagggatcttagggaac 636066
QY 4360 atgttgaataatgcccgtctcgatgatactgcttacaataaacttctgggttttcaggaajta 4419
DB 636067 atgttgaacaatgcaaggtttgaagatcgtctttcaacaactctcgggttcggctata 636126
QY 4420 ggaactttcttagctcaacaagaacactctcttcttcgagaagaattcagttactacagcgc 4479
DB 636127 ggaactttcttaggaaagaagtatctcgaaattctgactcattccactcatatggtcguga 636186
```





```
QY 1726 acaattccaggaatcac 1742
      |||||
Db 421 acaattccaggaatcac 437
      |||||

RESULT 10
ID AAA64715/c
XX AAA64715 standard; DNA; 395 BP.
AC AAA64715;
XX
XX
XX
XX 02-FEB-2001 (first entry)
DE
DE C. trachomatis LGV II clone 22-F8-91 coding sequence.
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial; ss.
XX
XX
XX
XX Chlamydia trachomatis.
XX
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US29012.
XX
XX 08-DEC-1999; 98US-0208277.
XX 08-APR-1999; 99US-0288594.
XX 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises
XX amino acid sequence encoded by polynucleotide sequence -
XX
XX Claim 1; Page 137; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumonia is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present sequence is a nucleic acid sequence
XX isolated in the present invention.
XX
XX Sequence 395 BP; 105 A; 82 C; 80 G; 128 T; 0 other;

Query Match 3.5%; Score 184.8; DB 21; Length 395;
Best Local Similarity 69.2%; Pred. No. 3.9e-39;
Matches 252; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 4209 taatcacaaacaggaattcaagccagatggacattcgatactatcgctcggt 4268
      |||||
Db 364 TCACTACAAAATGGAACGATTCAGCGCTGGAATTTGACTCTTATAGACAAATGGGC 305
      |||||
QY 4269 atacatacagggaataatcatttttttcggaacttatcttaggtcccaaaactcaat 4328
      |||||
```









**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 18:19:12 ; Search time 3705.76 Seconds  
(without alignments)  
13430.247 Million cell updates/sec

Title: US-09-454-684A-174  
Perfect score: 5265  
Sequence: 1 gcaatcatgaatggctgtc.....gtgctcgcatgatcttcttaa 5265

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues  
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_estl1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est33:\*  
26: gb\_est34:\*  
27: gb\_est35:\*  
28: gb\_est36:\*  
29: gb\_est37:\*  
30: gb\_est38:\*  
31: gb\_est39:\*  
32: gb\_est40:\*  
33: em\_estba:\*  
34: em\_estfun:\*  
35: em\_esthum1:\*  
36: em\_esthum2:\*  
37: em\_esthum3:\*  
38: em\_esthum4:\*  
39: em\_esthum5:\*  
40: em\_esthum6:\*  
41: em\_esthum7:\*  
42: em\_esthum8:\*  
43: em\_esthum9:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estin5:\*  
68: em\_estom1:\*  
69: em\_estom2:\*  
70: em\_estov1:\*  
71: em\_estov2:\*  
72: em\_estpl1:\*  
73: em\_estpl2:\*  
74: em\_estpl3:\*  
75: em\_estpl4:\*  
76: em\_estpl5:\*  
77: em\_estpl6:\*  
78: em\_estpl7:\*  
79: em\_estpl8:\*  
80: em\_estpl9:\*  
81: em\_estpl10:\*  
82: em\_estrol:\*  
83: em\_estrol2:\*  
84: em\_estrol3:\*  
85: em\_estrol4:\*  
86: em\_estrol5:\*  
87: em\_estrol6:\*  
88: em\_estrol7:\*  
89: em\_estrol8:\*  
90: em\_estrol9:\*  
91: em\_estrol10:\*  
92: em\_estrol11:\*  
93: em\_estrol12:\*  
94: em\_estrol13:\*  
95: em\_estrol14:\*  
96: em\_estrol15:\*  
97: em\_estrol16:\*  
98: em\_estrol17:\*  
99: em\_estrol18:\*  
100: em\_estrol19:\*  
101: em\_estrol20:\*  
102: gb\_est25:\*  
103: gb\_est26:\*  
104: gb\_est27:\*  
105: gb\_est28:\*  
106: gb\_est29:\*  
107: gb\_est30:\*  
108: gb\_est31:\*  
109: gb\_est32:\*  
110: gb\_est41:\*  
111: gb\_est42:\*  
112: gb\_est43:\*  
113: gb\_est44:\*  
114: gb\_est45:\*  
115: gb\_est46:\*  
116: gb\_est47:\*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
259: gb_est190:*
260: gb_est191:*
261: gb_est192:*
262: gb_est193:*
263: gb_est194:*
264: gb_est195:*
265: gb_est196:*
266: gb_est197:*
267: gb_est198:*
268: gb_est199:*
269: gb_est200:*
270: gb_est201:*
271: gb_est202:*
272: gb_est203:*
273: gb_est204:*
274: gb_est205:*
275: gb_est206:*
276: gb_est207:*
277: gb_est208:*
278: gb_est209:*
279: gb_est210:*
280: gb_est211:*
281: gb_est212:*
282: gb_est213:*
283: gb_est214:*
284: gb_est215:*
285: gb_est216:*
286: gb_est217:*
287: gb_est218:*
288: gb_est219:*
289: gb_est220:*
290: gb_est221:*
291: gb_est222:*
292: gb_est223:*
293: gb_est224:*
294: gb_est225:*
295: gb_est226:*
296: gb_est227:*
297: gb_est228:*
298: gb_est229:*
299: gb_est230:*
300: gb_est231:*
301: gb_est232:*
302: gb_est233:*
303: gb_est234:*
304: gb_est235:*
305: gb_est236:*
306: gb_est237:*
307: gb_est238:*
308: gb_est239:*
309: gb_est240:*
310: gb_est241:*
311: gb_est242:*
312: gb_est243:*
313: gb_est244:*
314: gb_est245:*
315: gb_est246:*
316: gb_est247:*
317: gb_est248:*
318: gb_est249:*
319: gb_est250:*
320: gb_est251:*
321: gb_est252:*
322: gb_est253:*
323: gb_est254:*
324: gb_est255:*
325: gb_est256:*
326: gb_est257:*
327: gb_est258:*
328: gb_est259:*
329: gb_est260:*
330: gb_est261:*
331: gb_est262:*
332: gb_est263:*
333: gb_est264:*
334: gb_est265:*
335: gb_est266:*
336: gb_est267:*
337: gb_est268:*
338: gb_est269:*
339: gb_est270:*
340: gb_est271:*
341: gb_est272:*
342: gb_est273:*
343: gb_est274:*
344: gb_est275:*
345: gb_est276:*
346: gb_est277:*
347: gb_est278:*
348: gb_est279:*
349: gb_est280:*
350: gb_est281:*
351: gb_est282:*
352: gb_est283:*
353: gb_est284:*
354: gb_est285:*
355: gb_est286:*
356: gb_est287:*
357: gb_est288:*
358: gb_est289:*
359: gb_est290:*
360: gb_est291:*
361: gb_est292:*
362: gb_est293:*
363: gb_est294:*
364: gb_est295:*
365: gb_est296:*
366: gb_est297:*
367: gb_est298:*
368: gb_est299:*
369: gb_est300:*
370: gb_est301:*
371: gb_est302:*
372: gb_est303:*
373: gb_est304:*
374: gb_est305:*
375: gb_est306:*
376: gb_est307:*
377: gb_est308:*
378: gb_est309:*
379: gb_est310:*
380: gb_est311:*
381: gb_est312:*
382: gb_est313:*
383: gb_est314:*
384: gb_est315:*
385: gb_est316:*
386: gb_est317:*
387: gb_est318:*
388: gb_est319:*
389: gb_est320:*
390: gb_est321:*
391: gb_est322:*
392: gb_est323:*
393: gb_est324:*
394: gb_est325:*
395: gb_est326:*
396: gb_est327:*
397: gb_est328:*
398: gb_est329:*
399: gb_est330:*
400: gb_est331:*
401: gb_est332:*
402: gb_est333:*
403: gb_est334:*
404: gb_est335:*
405: gb_est336:*
406: gb_est337:*
407: gb_est338:*
408: gb_est339:*
409: gb_est340:*
410: gb_est341:*
411: gb_est342:*
412: gb_est343:*
413: gb_est344:*
414: gb_est345:*
415: gb_est346:*
416: gb_est347:*
417: gb_est348:*
418: gb_est349:*
419: gb_est350:*
420: gb_est351:*
421: gb_est352:*
422: gb_est353:*
423: gb_est354:*
424: gb_est355:*
425: gb_est356:*
426: gb_est357:*
427: gb_est358:*
428: gb_est359:*
429: gb_est360:*
430: gb_est361:*
431: gb_est362:*
432: gb_est363:*
433: gb_est364:*
434: gb_est365:*
435: gb_est366:*
436: gb_est367:*
437: gb_est368:*
438: gb_est369:*
439: gb_est370:*
440: gb_est371:*
441: gb_est372:*
442: gb_est373:*
443: gb_est374:*
444: gb_est375:*
445: gb_est376:*
446: gb_est377:*
447: gb_est378:*
448: gb_est379:*
449: gb_est380:*
450: gb_est381:*
451: gb_est382:*
452: gb_est383:*
453: gb_est384:*
454: gb_est385:*
455: gb_est386:*
456: gb_est387:*
457: gb_est388:*
458: gb_est389:*
459: gb_est390:*
460: gb_est391:*
461: gb_est392:*
462: gb_est393:*
463: gb_est394:*
464: gb_est395:*
465: gb_est396:*
466: gb_est397:*
467: gb_est398:*
468: gb_est399:*
469: gb_est400:*
470: gb_est401:*
471: gb_est402:*
472: gb_est403:*
473: gb_est404:*
474: gb_est405:*
475: gb_est406:*
476: gb_est407:*
477: gb_est408:*
478: gb_est409:*
479: gb_est410:*
480: gb_est411:*
481: gb_est412:*
482: gb_est413:*
483: gb_est414:*
484: gb_est415:*
485: gb_est416:*
486: gb_est417:*
487: gb_est418:*
488: gb_est419:*
489: gb_est420:*
490: gb_est421:*
491: gb_est422:*
492: gb_est423:*
493: gb_est424:*
494: gb_est425:*
495: gb_est426:*
496: gb_est427:*
497: gb_est428:*
498: gb_est429:*
499: gb_est430:*
500: gb_est431:*
501: gb_est432:*
502: gb_est433:*
503: gb_est434:*
504: gb_est435:*
505: gb_est436:*
506: gb_est437:*
507: gb_est438:*
508: gb_est439:*
509: gb_est440:*
510: gb_est441:*
511: gb_est442:*
512: gb_est443:*
513: gb_est444:*
514: gb_est445:*
515: gb_est446:*
516: gb_est447:*
517: gb_est448:*
518: gb_est449:*
519: gb_est450:*
520: gb_est451:*
521: gb_est452:*
522: gb_est453:*
523: gb_est454:*
524: gb_est455:*
525: gb_est456:*
526: gb_est457:*
527: gb_est458:*
528: gb_est459:*
529: gb_est460:*
530: gb_est461:*
531: gb_est462:*
532: gb_est463:*
533: gb_est464:*
534: gb_est465:*
535: gb_est466:*
536: gb_est467:*
537: gb_est468:*
538: gb_est469:*
539: gb_est470:*
540: gb_est471:*
541: gb_est472:*
542: gb_est473:*
543: gb_est474:*
544: gb_est475:*
545: gb_est476:*
546: gb_est477:*
547: gb_est478:*
548: gb_est479:*
549: gb_est480:*
550: gb_est481:*
551: gb_est482:*
552: gb_est483:*
553: gb_est484:*
554: gb_est485:*
555: gb_est486:*
556: gb_est487:*
557: gb_est488:*
558: gb_est489:*
559: gb_est490:*
560: gb_est491:*
561: gb_est492:*
562: gb_est493:*
563: gb_est494:*
564: gb_est495:*
565: gb_est496:*
566: gb_est497:*
567: gb_est498:*
568: gb_est499:*
569: gb_est500:*
570: gb_est501:*
571: gb_est502:*
572: gb_est503:*
573: gb_est504:*
574: gb_est505:*
575: gb_est506:*
576: gb_est507:*
577: gb_est508:*
578: gb_est509:*
579: gb_est510:*
580: gb_est511:*
581: gb_est512:*
582: gb_est513:*
583: gb_est514:*
584: gb_est515:*
585: gb_est516:*
586: gb_est517:*
587: gb_est518:*
588: gb_est519:*
589: gb_est520:*
590: gb_est521:*
591: gb_est522:*
592: gb_est523:*
593: gb_est524:*
594: gb_est525:*
595: gb_est526:*
596: gb_est527:*
597: gb_est528:*
598: gb_est529:*
599: gb_est530:*
600: gb_est531:*
601: gb_est532:*
602: gb_est533:*
603: gb_est534:*
604: gb_est535:*
605: gb_est536:*
606: gb_est537:*
607: gb_est538:*
608: gb_est539:*
609: gb_est540:*
610: gb_est541:*
611: gb_est542:*
612: gb_est543:*
613: gb_est544:*
614: gb_est545:*
615: gb_est546:*
616: gb_est547:*
617: gb_est548:*
618: gb_est549:*
619: gb_est550:*
620: gb_est551:*
621: gb_est552:*
622: gb_est553:*
623: gb_est554:*
624: gb_est555:*
625: gb_est556:*
626: gb_est557:*
627: gb_est558:*
628: gb_est559:*
629: gb_est560:*
630: gb_est561:*
631: gb_est562:*
632: gb_est563:*
633: gb_est564:*
634: gb_est565:*
635: gb_est566:*
636: gb_est567:*
637: gb_est568:*
638: gb_est569:*
639: gb_est570:*
640: gb_est571:*
641: gb_est572:*
642: gb_est573:*
643: gb_est574:*
644: gb_est575:*
645: gb_est576:*
646: gb_est577:*
647: gb_est578:*
648: gb_est579:*
649: gb_est580:*
650: gb_est581:*
651: gb_est582:*
652: gb_est583:*
653: gb_est584:*
654: gb_est585:*
655: gb_est586:*
656: gb_est587:*
657: gb_est588:*
658: gb_est589:*
659: gb_est590:*
660: gb_est591:*
661: gb_est592:*
662: gb_est593:*
663: gb_est594:*
664: gb_est595:*
665: gb_est596:*
666: gb_est597:*
667: gb_est598:*
668: gb_est599:*
669: gb_est600:*
670: gb_est601:*
671: gb_est602:*
672: gb_est603:*
673: gb_est604:*
674: gb_est605:*
675: gb_est606:*
676: gb_est607:*
677: gb_est608:*
678: gb_est609:*
679: gb_est610:*
680: gb_est611:*
681: gb_est612:*
682: gb_est613:*
683: gb_est614:*
684: gb_est615:*
685: gb_est616:*
686: gb_est617:*
687: gb_est618:*
688: gb_est619:*
689: gb_est620:*
690: gb_est621:*
691: gb_est622:*
692: gb_est623:*
693: gb_est624:*
694: gb_est625:*
695: gb_est626:*
696: gb_est627:*
697: gb_est628:*
698: gb_est629:*
699: gb_est630:*
700: gb_est631:*
701: gb_est632:*
702: gb_est633:*
703: gb_est634:*
704: gb_est635:*
705: gb_est636:*
706: gb_est637:*
707: gb_est638:*
708: gb_est639:*
709: gb_est640:*
710: gb_est641:*
711: gb_est642:*
712: gb_est643:*
713: gb_est644:*
714: gb_est645:*
715: gb_est646:*
716: gb_est647:*
717: gb_est648:*
718: gb_est649:*
719: gb_est650:*
720: gb_est651:*
721: gb_est652:*
722: gb_est653:*
723: gb_est654:*
724: gb_est655:*
725: gb_est656:*
726: gb_est657:*
727: gb_est658:*
728: gb_est659:*
729: gb_est660:*
730: gb_est661:*
731: gb_est662:*
732: gb_est663:*
733: gb_est664:*
734: gb_est665:*
735: gb_est666:*
736: gb_est667:*
737: gb_est668:*
738: gb_est669:*
739: gb_est670:*
740: gb_est671:*
741: gb_est672:*
742: gb_est673:*
743: gb_est674:*
744: gb_est675:*
745: gb_est676:*
746: gb_est677:*
747: gb_est678:*
748: gb_est679:*
749: gb_est680:*
750: gb_est681:*
751: gb_est682:*
752: gb_est683:*
753: gb_est684:*
754: gb_est685:*
755: gb_est686:*
756: gb_est687:*
757: gb_est688:*
758: gb_est689:*
759: gb_est690:*
760: gb_est691:*
761: gb_est692:*
762: gb_est693:*
763: gb_est694:*
764: gb_est695:*
765: gb_est696:*
766: gb_est697:*
767: gb_est698:*
768: gb_est699:*
769: gb_est700:*
770: gb_est701:*
771: gb_est702:*
772: gb_est703:*
773: gb_est704:*
774: gb_est705:*
775: gb_est706:*
776: gb_est707:*
777: gb_est708:*
778: gb_est709:*
779: gb_est710:*
780: gb_est711:*
781: gb_est712:*
782: gb_est713:*
783: gb_est714:*
784: gb_est715:*
785: gb_est716:*
786: gb_est717:*
787: gb_est718:*
788: gb_est719:*
789: gb_est720:*
790: gb_est721:*
791: gb_est722:*
792: gb_est723:*
793: gb_est724:*
794: gb_est725:*
795: gb_est726:*
796: gb_est727:*
797: gb_est728:*
798: gb_est729:*
799: gb_est730:*
800: gb_est731:*
801: gb_est732:*
802: gb_est733:*
803: gb_est734:*
804: gb_est735:*
805: gb_est736:*
806: gb_est737:*
807: gb_est738:*
808: gb_est739:*
809: gb_est740:*
810: gb_est741:*
811: gb_est742:*
812: gb_est743:*
813: gb_est744:*
814: gb_est745:*
815: gb_est746:*
816: gb_est747:*
817: gb_est748:*
818: gb_est749:*
819: gb_est750:*
820: gb_est751:*
821: gb_est752:*
822: gb_est753:*
823: gb_est754:*
824: gb_est755:*
825: gb_est756:*
826: gb_est757:*
827: gb_est758:*
828: gb_est759:*
829: gb_est760:*
830: gb_est761:*
831: gb_est762:*
832: gb_est763:*
833: gb_est764:*
834: gb_est765:*
835: gb_est766:*
836: gb_est767:*
837: gb_est768:*
838: gb_est769:*
839: gb_est770:*
840: gb_est771:*
841: gb_est772:*
842: gb_est773:*
843: gb_est774:*
844: gb_est775:*
845: gb_est776:*
846: gb_est777:*
847: gb_est778:*
848: gb_est779:*
849: gb_est780:*
850: gb_est781:*
851: gb_est782:*
852: gb_est783:*
853: gb_est784:*
854: gb_est785:*
855: gb_est786:*
856: gb_est787:*
857: gb_est788:*
858: gb_est789:*
859: gb_est790:*
860: gb_est791:*
861: gb_est792:*
862: gb_est793:*
863: gb_est794:*
864: gb_est795:*
865: gb_est796:*
866: gb_est797:*
867: gb_est798:*
868: gb_est799:*
869: gb_est800:*
870: gb_est801:*
871: gb_est802:*
872: gb_est803:*
873: gb_est804:*
874: gb_est805:*
875: gb_est806:*
876: gb_est807:*
877: gb_est808:*
878: gb_est809:*
879: gb_est810:*
880: gb_est811:*
881: gb_est812:*
882: gb_est813:*
883: gb_est814:*
884: gb_est815:*
885: gb_est816:*
886: gb_est817:*
887: gb_est818:*
888: gb_est819:*
889: gb_est820:*
890: gb_est821:*
891: gb_est822:*
892: gb_est823:*
893: gb_est824:*
894: gb_est825:*
895: gb_est826:*
896: gb_est827:*
897: gb_est828:*
898: gb_est829:*
899: gb_est830:*
900: gb_est831:*
901: gb_est832:*
902: gb_est833:*
903: gb_est834:*
904: gb_est835:*
905: gb_est836:*
906: gb_est837:*
907: gb_est838:*
908: gb_est839:*
909: gb_est840:*
910: gb_est841:*
911: gb_est842:*
912: gb_est843:*
913: gb_est844:*
914: gb_est845:*
915: gb_est846:*
916: gb_est847:*
917: gb_est848:*
918: gb_est849:*
919: gb_est850:*
920: gb_est851:*
921: gb_est852:*
922: gb_est853:*
923: gb_est854:*
924: gb_est855:*
925: gb_est856:*
926: gb_est857:*
927: gb_est858:*
928: gb_est859:*
929: gb_est860:*
930: gb_est861:*
931: gb_est862:*
932: gb_est863:*
933: gb_est864:*
934: gb_est865:*
935: gb_est866:*
936: gb_est867:*
937: gb_est868:*
938: gb_est869:*
939: gb_est870:*
940: gb_est871:*
941: gb_est872:*
942: gb_est873:*
943: gb_est874:*
944: gb_est875:*
945: gb_est876:*
946: gb_est877:*
947: gb_est878:*
948: gb_est879:*
949: gb_est880:*
950: gb_est881:*
951: gb_est882:*
952: gb_est883:*
953: gb_est884:*
954: gb_est885:*
955: gb_est886:*
956: gb_est887:*
957: gb_est888:*
958: gb_est889:*
959: gb_est890:*
960: gb_est891:*
961: gb_est892:*
962: gb_est893:*
963: gb_est894:*
964: gb_est895:*
965: gb_est896:*
966: gb_est897:*
967: gb_est898:*
968: gb_est899:*
969: gb_est900:*
970: gb_est901:*
971: gb_est902:*
972: gb_est903:*
973: gb_est904:*
974: gb_est905:*
975: gb_est906:*
976: gb_est907:*
977: gb_est908:*
978: gb_est909:*
979: gb_est910:*
980: gb_est911:*
981: gb_est912:*
982: gb_est913:*
983: gb_est914:*
984: gb_est915:*
985: gb_est916:*
986: gb_est917:*
987: gb_est918:*
988: gb_est919:*
989: gb_est920:*
990: gb_est921:*
991: gb_est922:*
992: gb_est923:*
993: gb_est924:*
994: gb_est925:*
995: gb_est926:*
996: gb_est927:*
997: gb_est928:*
998: gb_est929:*
999: gb_est930:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	83.4	1.6	252	218	AF087168	AF087168 AF087168
2	63.6	1.2	946	15	AI069309	AI069309 mgae0006c
3	61.2	1.2	978	221	CNS040R1	AL269254 Tetraodon
4	61	1.2	885	152	BG368981	BG368981 HVSMei002
5	61	1.2	1003	221	CNS040U2	AL303092 Tetraodon
6	60.6	1.2	536	222	FR0036552	AL124061 Fugu rubr
7	60.2	1.1	906	220	CNS02ILL	AL199074 Tetraodon
8	60.2	1.1	962	222	CNS054Y6	AL321351 Tetraodon
9	58.6	1.1	808	220	CNS02NGR	AL205380 Tetraodon
10	58.4	1.1	269	247	AZ662861	AZ662861 IM0542M10
11	57.8	1.1	615	247	AZ661546	AZ661546 IM0540M03
12	57.8	1.1	712	247	AZ661544	AZ661544 IM0540M01
13	57.8	1.1	2199	139	BE741896	BE741896 601595356
14	57.6	1.1	770	139	BE744403	BE744403 601576651
15	57.2	1.1	667	221	CNS03YBK	AL266285 Tetraodon
16	57	1.1	415	250	AZ839702	AZ839702 2M0135N24
17	56.6	1.1	1587	152	BG338956	BG338956 602436760
18	56	1.1	987	132	AK013560	AK013560 Mus muscu
19	55.8	1.1	727	251	AZ870651	AZ870651 2M0183011
20	55.6	1.1	562	240	AZ251135	AZ251135 RPI-23-5
21	55.6	1.1	700	105	AL508461	AL508461 AL508461
22	55.4	1.1	405	1	AA065249	AA065249 al2501r T
23	55.4	1.1	439	188	R98242	R98242 YQ78e03.r1
24	55.2	1.0	455	155	BG561862	BG561862 EcEtree03
25	55	1.0	226	250	AZ837781	AZ837781 2M0133011
26	54.8	1.0	489	146	BF287438	BF287438 EST452029
27	54.8	1.0	925	140	BE778075	BE778075 601463186
28	54.6	1.0	966	219	CNS004W2	AL055504 Drosophil
29	54.6	1.0	980	219	CNS01PHD	AL154817 Anopheles
30	54.4	1.0	264	140	BE798211	BE798211 601585978
31	54.4	1.0	351	222	FR0004775	288558 F.rubripes
32	54.4	1.0	555	241	AZ281461	AZ281461 RPI-23-1
33	54.4	1.0	619	222	FR0013713	AL004959 F.rubripe
34	54.4	1.0	993	220	CNS03AOO	AL235545 Tetraodon
35	54.2	1.0	599	163	BE115577	BE115577 UT-R-BJ1-
36	54	1.0	540	116	AW464269	AW464269 BP230015A
37	54	1.0	648	18	AI257238	AI257238 LP05546.5
38	54	1.0	987	122	AW940443	AW940443 GH12043.3
39	53.8	1.0	250	115	AW409026	AW409026 fbl_d10 F
40	53.8	1.0	935	220	CNS033D4	AL225985 Tetraodon
41	53.6	1.0	162	222	FR0009521	AL000805 F.rubripe
42	53.4	1.0	354	116	AW486593	AW486593 75069 MAR
43	53.4	1.0	477	107	AU084733	AU084733 AU084733
44	53.4	1.0	680	19	AI388163	AI388163 GH18971.5
45	53.4	1.0	691	246	AZ627417	AZ627417 IM0469P07

## ALIGNMENTS

RESULT	1
AF087168	
LOCUS	AF087168 252 bp DNA
DEFINITION	AF087168 Chlamydia trachomatis L2 GSS 18-APR-2000
	clone 205 similar to type 1-related protein phosphatase, DNA
ACCESSION	AF087168
VERSION	AF087168.1 GI:7593836
KEYWORDS	GSS
SOURCE	Chlamydia trachomatis.
ORGANISM	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE	1 (bases 1 to 252)
AUTHORS	Wang, L., Steenburg, S.D., Zheng, Y. and Larsen, S.H.
TITLE	Gene identification of Chlamydia trachomatis by random DNA sequencing
JOURNAL	Unpublished (1998)

## COMMENT

Contact: Wang L  
Department of Microbiology & Immunology  
Indiana University School of Medicine  
635 Barnhill Drive; MS 255, Indianapolis, IN 46202, USA  
Class: Shotgun.

## FEATURES

Location/Qualifiers  
1..252  
/organism="Chlamydia trachomatis"  
/strain="L2"  
/db\_xref="taxon:813"  
/clone="205"  
/clone\_lib="Chlamydia trachomatis L2"  
/note="Isolate=434B"  
77 a 54 c 54 g 67 t  
ORIGIN

Query Match 1.6%; Score 83.4; DB 218; Length 252;  
Best Local Similarity 87.9%; Pred. No. 3e-11;  
Matches 138; Conservative 0; Mismatches 11; Indels 8; Gaps 4;

QY 4460 aattcagttactacagccgaaacttcagttgcccacatg----ccaaacctagacaa 4515  
|||||  
Db 93 AATTCAGTTACTACACCCGCGAACTTCAAGTTGCCATCAGGATGCCAAACCTAGACAA 152  
QY 4516 ga-ttttatctcagagctgcatttagtaagatagtggggaaaccaaagccatcaaaa 4574  
|||||  
Db 153 GATTTTATCCTAGGAGCTGCATTTAGTAGATGGTGGGAAACCAACCAAGCCATCAANA 212  
QY 4575 atgcataattacttcataaggctcctgagtactct 4611  
|||||  
Db 213 AATGCAT-ATTACTTCCATA--GGCTCTGAGTACTCT 246

## RESULT 2

AI069309 946 bp mRNA EST 09-DEC-1999  
LOCUS mgae0006cc09f Magnaporthe grisea Appressorium Stage cDNA Library  
DEFINITION Magnaporthe grisea cDNA clone mgae0006cc09f 5', mRNA sequence.  
ACCESSION AI069309  
VERSION AI069309.1 GI:3392284  
KEYWORDS EST.  
SOURCE Magnaporthe grisea.  
ORGANISM Magnaporthe grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.  
REFERENCE 1 (bases 1 to 946)  
AUTHORS Choi, W., Fang, E., Sasinowski, M., Wing, R. and Dean, R.A.  
TITLE Expressed sequence characterization during appressorium formation in rice blast fungus, Magnaporthe grisea  
JOURNAL Unpublished (1998)  
COMMENT Contact: Dean, R.A.  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson University, Clemson, SC 29634, USA  
Tel: 864 656 5737  
Fax: 864 656 4293  
Email: rdean@clemson.edu  
Seq primer: T3 primer (AATTAAACCTCACTAAAGGG)  
High quality sequence stop: 403.  
Location/Qualifiers  
1..946  
/organism="Magnaporthe grisea"  
/strain="70-15"  
/db\_xref="taxon:148305"  
/clone="mgae0006cc09f"  
/clone\_lib="Magnaporthe grisea Appressorium Stage cDNA Library"  
/dev\_stage="Germinated conidia on appressorium-inductive surface"  
/note="Vector: pBlueScriptII SK(+) Vector; Site1: EcoRI; Site2: XhoI; The appressorium formation-specific cDNA library was constructed from conidia germinated for 5-8

## FEATURES

source  
1..946  
/organism="Magnaporthe grisea"  
/strain="70-15"  
/db\_xref="taxon:148305"  
/clone="mgae0006cc09f"  
/clone\_lib="Magnaporthe grisea Appressorium Stage cDNA Library"  
/dev\_stage="Germinated conidia on appressorium-inductive surface"  
/note="Vector: pBlueScriptII SK(+) Vector; Site1: EcoRI; Site2: XhoI; The appressorium formation-specific cDNA library was constructed from conidia germinated for 5-8



















;; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and  
;; TITLE OF INVENTION: Primers Therefor  
;; FILE REFERENCE: 0760-0241p  
;; CURRENT APPLICATION NUMBER: US/09/043.303  
;; CURRENT FILING DATE: 1998-05-18  
;; EARLIER APPLICATION NUMBER: PCT/JP96/01999  
;; EARLIER FILING DATE: 1996-07-18  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 7  
;; LENGTH: 203  
;; TYPE: DNA  
;; ORGANISM: p-2093 plasmid  
US-09-043-303-7

Query Match 1.1%; Score 58.6; DB 4; Length 203;  
Best Local Similarity 64.2%; Pred. No. 1.3e-06;  
Matches 88; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
Qy 2392 aaagcaaaactctccacagaggtcgcaaacagcttcagcacctgtgcagctgct 2451  
Db 12 acagcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagca 71  
Qy 2452 gcttcctacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2511  
Db 72 gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagca 131  
Qy 2512 ggtgtagtaggagc 2528  
Db 132 gcagcagcagcagcagc 148

RESULT 3  
5273901-6  
; Patent No. 5273901  
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;  
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
; TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS  
; SPOOROITE 21.5 KB ANTIGEN, AC-6B  
; NUMBER OF SEQUENCES: 11  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/581.693  
; FILING DATE: 12-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 215,162  
; FILING DATE: 05-JUL-1988  
; APPLICATION NUMBER: 746,520  
; FILING DATE: 19-JUN-1985  
; APPLICATION NUMBER: 627,811  
; FILING DATE: 05-JUL-1984  
; SEQ ID NO:6:  
; LENGTH:543  
5273901-6

Query Match 1.1%; Score 58; DB 6; Length 543;  
Best Local Similarity 61.0%; Pred. No. 3.5e-06;  
Matches 94; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
Qy 2392 aaagcaaaactctccacagaggtcgcaaacagcttcagcacctgtgcagctgct 2451  
Db 154 acagcagctgcaacagcagcagcagcagcagcagcagcagcagcagcagcagca 213  
Qy 2452 gcttcctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2511  
Db 214 gcagcagcagctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 273  
Qy 2512 ggtgtagtaggagcagctatctatgagaaaag 2545  
Db 274 acagccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 307

RESULT 4  
US-08-469-802B-2  
; Sequence 2, Application US/08469802B  
; Patent No. 5741645  
; GENERAL INFORMATION:  
; APPLICANT: Orr, Harry T.  
; APPLICANT: Rarum, Laura P.W.  
; APPLICANT: Chung, Ming-yi  
; APPLICANT: Zoghbi, Huda Y.  
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
; Patent No. 5741645  
; TITLE OF INVENTION: Type 1 and Method for Diagnosis  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Muetting, Raasch, Gebhardt & Schwappach, P.A.  
; STREET: 119 No. 5741645th Fourth Street, Suite 203  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/469,802B  
; APPLICATION NUMBER: US/08/469,802B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Muetting, Ann M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 110.00030101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1225  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 195 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-469-802B-2

Query Match 1.1%; Score 57.2; DB 1; Length 195;  
Best Local Similarity 57.1%; Pred. No. 3.2e-06;  
Matches 104; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
Qy 2394 agcaaaactctccacagaggtcgcaaacagcttcagcacctgtgcagctgctgc 2453  
Db 14 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 73  
Qy 2454 ttcctacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2513  
Db 74 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 133  
Qy 2514 tgtagtaggagcagctatctatggagaaaagttacattctctcaatgtagcggacttg 2573  
Db 134 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 193  
Qy 2574 tc 2575  
Db 194 TC 195  
RESULT 5  
US-08-267-803B-2  
; Sequence 2, Application US/08267803B  
; Patent No. 5834183  
; GENERAL INFORMATION:

```

RESULT      6
US-08-469-802B-3
; Sequence 3, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Ori, Harry T.
; APPLICANT: Ratum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645

```

```

Query Match      1.1%; Score 57.2; DB 1; Length 234;
Best Local Similarity 57.1%; Pred. No. 3.5e-06;
Matches 104; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 2394 agcaactctctatccacagaggagtgcgaacaacagcttcagcacctgtcgtcgacctgctgc 2453
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 53 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 112
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2454 ttccctaacaagcagccgcagcagcgccaccatcatctccagacaacaccaacttattcagg 2513
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 113 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 172
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2514 tgttagtgaggaggtctatctgtggagaagaagtttacattctctcaatgtacgggaacttg 2573
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 173 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 232
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2574 tc 2575
    ||
DB 233 TC 234

RESULT 7
US-08-267-803B-3
; Sequence 3, Application US/08267803B
; Patent No. 5834183
; GENERAL INFORMATION:
; APPLICANT: Ori, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5834183
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mueting, Raasch, Gebhardt & Schwappach, P.A.
; STREET: P.O. Box 581415
; CITY: Minneapolis

```

```
Query Match      1.1%; Score 56.6; DB 1; Length 154;
Best Local Similarity 63.7%; Pred. No. 4e-06;
Matches 86; Conservative 0; Mismatches 49; Indels 0; Caps 0;
```

Qy	2394	agcaaaactcctcatccaaagagagtcgcaaacacacagcttccagcaactctctcagctgctgc	24153
Db	3	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	62
Qy	2454	tccctacaagcagccgcgagccgcgaccatctatctccagcaacacaaacttattcagg	2513
Db	63	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	122
Qy	2514	tgtagtaggaggagc	2528
Db	123	AGCAGCAGCAGCAGC	137

```

RESULT          9
US-08-267-803B-6
; Sequence 6, Application US/08267803B
; Patent No. 5834183
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5834183
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muelling, Raasch, Gebhardt & Schwappach, P.A.
; STREET: P.O. Box 58415
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55458-1415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/267, 803B
; FILING DATE: 28-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McCormack, Myra H.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 110.000301120

```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-305-1217  
TELEFAX: 612-305-1228  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 154 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-267-803B-6

Query Match 1.1%; Score 56.6; DB 2; Length 154;  
Best Local Similarity 63.7%; Pred. No. 4e-06;  
Matches 86; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 2394 agcaaaactctctacagaggagtcgcaacaacagcttccagcacctgtcgtcagctgtgc 2453  
Db 3 AGC 62  
QY 2454 ttccctacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2513  
Db 63 AGC 122  
QY 2514 tgtagtaggaggagc 2528  
Db 123 AGC 177

RESULT 10  
US-09-043-303-17  
Sequence 17, Application US/09043303  
Patent No. 6251589  
GENERAL INFORMATION:  
APPLICANT: TSUJI, Shoji  
TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and  
TITLE OF INVENTION: Primers Therefor  
FILE REFERENCE: 0760-0241P  
CURRENT APPLICATION NUMBER: US/09/043,303  
CURRENT FILING DATE: 1998-05-18  
EARLIER APPLICATION NUMBER: PCT/JP96/01999  
EARLIER FILING DATE: 1996-07-18  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 17  
LENGTH: 165  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Probe  
US-09-043-303-17

Query Match 1.1%; Score 56.6; DB 4; Length 165;  
Best Local Similarity 63.7%; Pred. No. 4.2e-06;  
Matches 86; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 2394 agcaaaactctctacagaggagtcgcaacaacagcttccagcacctgtcgtcagctgtgc 2453  
Db 2 agc 61  
QY 2454 ttccctacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2513  
Db 62 agc 121  
QY 2514 tgtagtaggaggagc 2528  
Db 122 agc 136

RESULT 11

US-08-469-802B-4  
Sequence 4, Application US/08469802B  
Patent No. 5741645  
GENERAL INFORMATION:  
APPLICANT: Orr, Harry T.  
APPLICANT: Ranum, Laura P.W.  
APPLICANT: Chung, Ming-yi  
APPLICANT: Zoghbi, Huda Y.  
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
Patent No. 5741645  
TITLE OF INVENTION: Type 1 and Method for Diagnosis  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Muetting, Raasch, Gebhardt & Schwappach, P.A.  
STREET: 119 No. 5741645th Fourth Street, Suite 203  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,802B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Muetting, Ann M.  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 110.00030101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-305-1217  
TELEFAX: 612-305-1225  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 168 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-469-802B-4

Query Match 1.1%; Score 56.6; DB 1; Length 168;  
Best Local Similarity 63.7%; Pred. No. 4.2e-06;  
Matches 86; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 2394 agcaaaactctctacagaggagtcgcaacaacagcttccagcacctgtcgtcagctgtgc 2453  
Db 2 AGC 61  
QY 2454 ttccctacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2513  
Db 62 AGC 121  
QY 2514 tgtagtaggaggagc 2528  
Db 122 AGC 136  
RESULT 12  
US-08-267-803B-4  
Sequence 4, Application US/08267803B  
Patent No. 5834183  
GENERAL INFORMATION:  
APPLICANT: Orr, Harry T.  
APPLICANT: Ranum, Laura P.W.  
APPLICANT: Chung, Ming-yi  
APPLICANT: Zoghbi, Huda Y.  
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
Patent No. 5834183

```

APPLICANT: ZOGNDI, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267,803B
FILING DATE: 28-Jun-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.

```





**THIS PAGE BLANK (USPFO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2001, 08:34:51 ; Search time 30.88 Seconds

(without alignments)  
3439.348 Million cell updates/sec

Title: US-09-454-684A-180

Perfect score: 8942  
Sequence: 1 MKWLSATVFAVPAVPSVSGF.....IDVGWYTLQMTSCGARMIF 1752

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0601:\*

1:	/SID8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2:	/SID8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3:	/SID8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4:	/SID8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5:	/SID8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6:	/SID8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7:	/SID8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8:	/SID8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9:	/SID8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10:	/SID8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11:	/SID8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12:	/SID8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13:	/SID8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14:	/SID8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15:	/SID8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16:	/SID8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17:	/SID8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18:	/SID8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19:	/SID8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20:	/SID8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21:	/SID8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22:	/SID8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	8942	100.0	1752	21	AA13637
2	5858	65.5	1194	20	AA137825
3	3592	40.2	1776	21	AA13636
4	2703	30.2	948	21	AA13643
5	2621	29.3	892	20	AA137822
6	2568.5	28.7	568	20	AA137828
7	2474.5	27.7	1146	20	AA135162
8	1272.5	14.2	469	21	AA195547
9	1002	11.2	915	20	AA137823
10	943.5	10.6	821	21	AA13644
11	748.5	8.4	1617	20	AA135617

12	742.5	8.3	1132	20	AA135048	Chlamydia pneumoniae
13	685.5	7.7	1530	21	AA13635	C. trachomatis pmp
14	665.5	7.4	930	20	AA135052	Chlamydia pneumoniae
15	660.5	7.4	930	21	AA190240	Chlamydia antigen
16	655.5	7.3	914	20	AA188429	Chlamydia pneumoniae
17	655.5	7.3	928	20	AA188418	Chlamydia pneumoniae
18	652.5	7.3	930	20	AA188424	Chlamydia pneumoniae
19	636	7.1	609	20	AA135161	Chlamydia pneumoniae
20	619.5	6.9	928	21	AA190237	Chlamydia antigen
21	616	6.9	885	21	AA190238	Mature Chlamydia a
22	567	6.3	871	21	AA195550	Chlamydia pneumoniae
23	560	6.3	928	20	AA188421	Chlamydia pneumoniae
24	559	6.3	928	21	AA194327	Chlamydia pneumoniae
25	555.5	6.2	925	21	AA199843	Chlamydia pneumoniae
26	555.5	6.2	936	21	AA199842	Chlamydia pneumoniae
27	555	6.2	375	21	AA195549	Chlamydia pneumoniae
28	551.5	6.2	931	21	AA192715	Chlamydia pneumoniae
29	529.5	5.9	922	20	AA188419	Chlamydia pneumoniae
30	525.5	5.9	922	21	AA195548	Chlamydia pneumoniae
31	524.5	5.9	918	21	AA169369	Amino acid sequenc
32	523.5	5.9	918	20	AA188422	Chlamydia pneumoniae
33	522.5	5.8	922	20	AA134597	Chlamydia pneumoniae
34	517.5	5.8	949	20	AA135060	Chlamydia pneumoniae
35	516.5	5.8	928	20	AA188417	Chlamydia pneumoniae
36	516.5	5.8	928	21	AA190236	Chlamydia antigen
37	506	5.7	841	21	AA192818	C. pneumoniae CPN1
38	504	5.6	841	20	AA188420	Chlamydia pneumoniae
39	498	5.6	1013	20	AA16738	C. trachomatis F s
40	496.5	5.6	643	20	AA135056	Chlamydia pneumoniae
41	492	5.5	928	20	AA188423	Chlamydia pneumoniae
42	491	5.5	1013	20	AA16737	C. trachomatis B s
43	490	5.5	928	21	AA190239	Chlamydia antigen
44	489	5.5	962	21	AA13633	C. trachomatis pmp
45	485	5.4	1006	21	AA13639	C. trachomatis pmp

#### ALIGNMENTS

RESULT 1	
ID	AA13637 standard; Protein: 1752 AA.
AC	AA13637:
XX	
XX	02-FEB-2001 (first entry)
XX	
XX	C. trachomatis pmp gene protein.
DE	
XX	
XX	Chlamydia infection; sexually transmitted disease;
KW	pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW	trachoma; blindness; acute respiratory tract infection;
KW	atherosclerosis; coronary heart disease; antibacterial.
XX	
OS	Chlamydia trachomatis.
XX	
PN	WO200034483-A2.
XX	
PD	15-JUN-2000.
XX	
PF	08-DEC-1999; 99WO-US29012.
XX	
PR	08-DEC-1998; 98US-0208277.
PR	08-APR-1999; 99US-0288594.
PR	01-OCT-1999; 99US-0410568.
XX	22-OCT-1999; 99US-0426571.
XX	
PA	(COR-) CORTXA CORP.
XX	
PI	Probst P, Bhatia A, Skelky YAW, Fling SP, Jen S, Stromberg EJ;
XX	WPI; 2000-431303/37.
XX	

PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -

Claim 2: Pages 194-198; 256pp; English.

XX The present invention relates to new nucleic acid sequences and the  
XX proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydia are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a protein isolated in the  
CC present invention.

XX Sequence 1752 AA;

Query Match 100.0%; Score 8942; DB 21; Length 1752;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MWLISATFAAVALPVSQFCPEPEPKELNFSRYETSSSTFTTIGGAEIVTSGNASF 60  
DB 1 mwlisatfaavilpvsqfcfpepkelnfsretssstfttcitigegaeylvsgnaaf 60  
QY 61 TFFNTIPTDTTTPNNSSSSSGSETASVSEDSSTTTTTPDKGGAIFYNAHSGVLEFMT 120  
DB 61 tffntiptdtttptnnsnsssgsetasvseedssttttppdkyggaifynahgvislmt 120  
QY 121 RSGTEGSLTSEIKMTGEGGAIFSGEELLPTDLTSLTTONNLSOLSGAIFGSGTSLSG 180  
DB 121 rsgtesltselkmtgeggailsgeellftdltsltlqnnlsqsgaifgsgtstlsq 180  
QY 181 ITRKATSCNSAEVPAPVKKPTEPKAQTASETSGSSSSGNDVSVPSSSRAEPAANLOS 240  
DB 181 itkatcsnsaevpapvkkptepkaqtasetsgsssgndvsvspssraepaanaals 240  
QY 241 HETCATATPAAGDTETSTPSPSHKPGSGAIYAKGDIITADSOEVLFSINKATKDGGAIFA 300  
DB 241 hfcatatpaagdtetstpsphkpgsgaiyakgdilitadsoevlfsinkatkdggaifa 300  
QY 301 EKQVSENTTSLKVONNGAEKGAIFYAKGDISIOSSKOSLFMSNYSKOGGALLYVEGCI 360  
DB 301 ekqvstenltslkvongaeekgaifyakgdlsioSSKqslfmsnyskqggallvegqi 360  
QY 361 NFQDLEIRIKYNAKGTFFETKKTITLPLSLKAQASAGNADMASSPSGSGATVSDSGS 420  
DB 361 nfgdleirikyngakgtfekkkitlpslkqasagnadawasspsgsgatlvssgds 420  
QY 421 SSGSDSDSETVPPVYAKGGGLYTDKNLSTNTITGIETIANNKATDVGGAYVKGTLTCEN 480  
DB 421 ssgsdsetvpyakggglytdknlstntitgiETIANnkATdvGGayvkgtltcen 480  
QY 481 SHRLQFLKNSDKOGGIVGENDITLNLTKGTLPOENAKKEGGGLFTKGTOKALTMMG 540  
DB 481 shrlqflknsdkggglygedniltlnltgkltlfgentakegggllfkgtokaltmng 540  
QY 541 LDSFCLINNTSEKHGGAFTTKEISQTYTSDVETIGITPVHGEITYTGKSTGAGNGCV 600  
DB 541 ldsfclinntsekhggafTtkeIsqTytsdvETIGITpvHgeITyTGKstGAGngcv 600  
QY 601 CTIKRALSNLQSTISISGNSAENGGAHTCCPDSPPTADTAEOGPAASAATSTPKSAVST 660  
DB 601 ctkralnslqstisigNSAaenggahTccpdspPTAdTAEOGpAAsAATstPKsAvst 660  
QY 661 ALSTPSSIVSSLTLAASSQASPATSNKETQPNADDTLLLDYVVDTTISKRTAKKGG 720

DB 661 alstpsstvssltllaassqaspatSNketQPNadDTLLldYvvDTtIsKrtAKKgg 720  
QY 721 IYAKKAKMRIDOLNISENSATEIGGICKESELEDAVLSTVENTLKGKGGGHAAT 780  
DB 721 iyakkakmrIdolnIsenSateIggicKESeLEdAVLstvEntlKgKgGghAat 780  
QY 781 VNISNLKSGPESNNKANSSSTGCVATTASAPAAAALQAAAAAPSPATPYSGVGG 840  
DB 781 vnislksgfEsnnkanSSstGvatTAsAPAAAalQAAAApSPatPySGvGg 840  
QY 841 AIYGEVTFESQSGCTPOFGNQAIDNNPQSSILNVGGAIYVKTSLTSSDAGTIFPS 900  
DB 841 aiylgevtfEsqsgctPoFGnQAIdnnPqSSilnvGgAIyVktSlTssDagTifPs 900  
QY 901 GNSVSTGKSQTTGOIAGAIYSPPTVLNCPAYFNNNTAIAIPKTSSEGGSGNSIKDTI 960  
DB 901 gnsvstgksqTtgoIAGAIySPptVLNcPAYfNNNTAIAIpKtsSEgSGnsIKdtI 960  
QY 961 GCATAGTAITLGSVRSFGNFDLGAATGTLANANTPSATSGSQNSITETITLNGSPTF 1020  
DB 961 gcatagtaItlGSvRSfgnFDlGAATgtLANantPSatSGsqNsITetITlNgSPtF 1020  
QY 1021 ERNOANKRGAIVSPSYSTIGNNTENONTSTHDGSAIYFTKATIESLGSVLFTGNMVA 1080  
DB 1021 ernOankrGAivSPsyStIGnntEnONTstHDgSAIyftKATIESlgSVlftGnmvA 1080  
QY 1081 TQASSATSGCNTTANYGAATFGDPCTTOSQTDALITTLAASSGNTTFNNSLONNQDPT 1140  
DB 1081 tqassatsgcnTtanYGAatFGdpCTtOSqTdAlITtLlaSSgntTFnNsLoNNqDpt 1140  
QY 1141 PASKFCSIGYVYKLSLQAAGKTISSFDCVHNTSTKKTGSTQNYVETLIDINKENSNPTYG 1200  
DB 1141 paskfcsIgyvYkLsLQAAGkTIsSFdcVhntStKktGstQnyVetLIIdINKeNSnPtYg 1200  
QY 1201 TIVFSESELHENSXYIPQNALIHNGTLVLKEKTELHVYSEPOKGSGLIMEPGAVLSNOMI 1260  
DB 1201 tivfseSelHensxyIpqNaLIhngTlVLkEkTeLHVySePoKGSglImEPgavLSnOmI 1260  
QY 1261 ANGALAINGLITDLSMGTPQAGEIFSPPELRIVATTSASGGSGVSSSIPTPNKRISAA 1320  
DB 1261 angalainGLITdLSmgTpQageIfsPpELrIVatTSaSGSGvSSsIPtpNkrISaa 1320  
QY 1321 VSGSAATTPPMSENVFLTGDLITLDPGNFYQNPMLGSLDVLPIKIPNTSDVQVYD 1380  
DB 1321 vsgsaattpPMseNVflTGdlITlDpGnfyQnpMLgSLdVlPIkIPntSDvQvYd 1380  
QY 1381 LFTLSGLPPOKGYMWTLDSPQGTGLQARWTFDVRBWVYIPRNHRYANSIIGSQNS 1440  
DB 1381 lftlsGLpPOkGYmWTldSPqGTglQARwTFdVRBWvYIPrNHrYANSiIGsqNs 1440  
QY 1441 MIVVKOGLINMMLNARFDDIAYNFVWVGVTFLAQOGTPLSEEPSYYSRGTSVADIAK 1500  
DB 1441 mivvkogLIInmMLnARfDDIAynFvWvgVtFlAQOGtPLsEePSyYSrGtSVAdIAk 1500  
QY 1501 PRODFLGAAFSKIVGKTAIKKMHNTYFKGSEYSTQASVYGGKFLYFLLNKOHGALEF 1560  
DB 1501 prOdfLGAafSKiVGkTAIkKmhNTyfkGseYstQAsvYgGkFlYfLlnKohGaLEf 1560  
QY 1561 LIOGVVSYGKIKHDTTTLVPSIHERNKGWEDLGMFLADIRISMDLKEPKDSKRTTYVG 1620  
DB 1561 liOgvvSYgKikHdTtTLvPSiHERnKGwEDlGMfLADIRISMDlKEpKdsKrtTyVg 1620  
QY 1621 ELEYSIRKOFPEIYDPRHFDCAVRYNLSTLPVGAIVEGAILMNCNIMLYNKLALAYMS 1680  
DB 1621 eleySiRkOfPEiYdPrHfDcAVryNLstLPvGAiVEgAILmNCnIMlyNklAlAYms 1680  
QY 1681 IYNNPNVCYRIVLSSNDAGOVICGVPRTTSARAENVSTQLYLGFWTLTIGNYTIDVGMYTL 1740  
DB 1681 iYnnpNvcYrIVlSSNdAGovICgVpRTtsARaENVstQlYlGfWtLTiGnyTiDvGmYtl 1740  
QY 1741 SQMTSGARMIIF 1752

Db	1741	sqmtscgarmif	1752		
	RESULT	2			
	AAAY37825				
ID	AAAY37825	standard; Protein; 1194	AA.		
XX	AAAY37825;				
XX	07-OCT-1999	(first entry)			
XX	Chlamydia trachomatis	cellular envelope protein.			
XX	Vaccine; eye disease; conventional trachoma; nonendemic trachoma; paratrachoma; inclusion conjunctivitis; genital disease; perinephatitis; nongonococcal urethritis; epididymitis; cervicitis; salpingitis; Bartholinitis; pneumopathy; venereal lymphogranulomatosis.				
OS	Chlamydia trachomatis.				
XX	WO9928475-A2.				
XX	10-JUN-1999.				
XX	27-NOV-1998;	98WO-IB01939.			
XX	04-NOV-1998;	98US-0107077.			
XX	28-NOV-1997;	97FR-0015041.			
XX	17-DEC-1997;	97FR-0016034.			
XX	(GEST ) GENSET.				
XX	Griffais R;				
XX	WPI; 1999-371125/31.				
XX	Genome sequence of Chlamydia trachomatis				
XX	Disclosure; Page 1395-1396; 1755pp; English.				
XX	AAAY36754-Y37949	are encoded by open reading frames (ORFs) of the genome of Chlamydia trachomatis (see AA01425). The polypeptides can be used as vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences can also be used to control growth of the microorganism. Chlamydia trachomatis is responsible for a large number of diseases, e.g. eye diseases such as conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion conjunctivitis; genital diseases such as nongonococcal urethritis, epididymitis, cervicitis, salpingitis, perinephatitis, Bartholinitis; pneumopathy in breast feeding infants; and venereal lymphogranulomatosis. The polypeptides of the invention may be of use in treating these diseases.			
XX	Sequence	1194	AA;		
QY	Query Match	65.5%;	Score 5858;	DB 20; Length 1194;	
DB	Best Local Similarity	97.3%;	Pred. No. 1.7e-289;		
DB	Matches 1147;	Conservative	12;	Mismatches 16; Indels 4; Gaps 2;	
QY	562	KEISQTTSDVETIPGTPVHGTVITGNKSTGNGGVCVKRRLALSNLQISISGNSAA	621		
DB	16	keisqtytsdvetiipgitpvhgetvitgkntgnggvcvkrlalsnlqsisisgnsaa	75		
QY	622	ENGGAHTCPDSPFTADTAEPAAASATSPKAPV-STALSTPSSSTVSSLTLAASS	680		
DB	76	engggaytcpcdsftadtaeapaaasatstpcsdpvastaltspssstvasstllaass	135		
QY	681	QASPATSNKETQDPNADTDLIDVVDVTTISKNTAKKGGGIYAKKAKMSRIDQLNISNS	740		
DB	136	qaspatstnektqdpnadtdlldyvvdttskntakkg99gyakkakmsridqlnisns	195		
QY	741	ATEITGGICCKESLDALVSLSTENLVGKEGGGLHAKTVYNISNLKSGFSFNKANS	800		



Db 1071 gdvtkygaalfiglasngsqtdnlplkliasggnicfrnneyrptssdtgtstfcsiaq 1130  
QY 1151 YVKLSQAARKGTISFDCVHTSTKKTGSTQNVYETLDINKENS---NPYTGTVFSS 1206  
Db 1131 dvkltmqakgktsifdairtktgtatqataydldinksedsetvsafgtgtilfss 1190  
QY 1207 ELHENSYPQNALHNGTLVLEKTELHVVSFEQKESKLIMEPGAVLSNQNIANGALA 1266  
Db 1191 elhenksyipqnvvlhsgsilvlpntelhlvisfeqegsalvntpgvslnqtvadgalv 1250  
QY 1267 INGLTTDLSSM---GTPQAGEIFSPPELRIVATTSASGG----- 1303  
Db 1251 inmtldlssveknglae-gniftppehrlidtttsgsggtptdsesngnsdtkteqnn 1309  
QY 1304 -----SGVSSSIP-----TNPKRISAAPVSGSAATTP-----TMSENKVFUTGDLTLI 1346  
Db 1310 ndasngesangssspavaaahstrtrnfaaataatpttptatttsnqylggekili 1369  
QY 1347 DPNGNYQNPMLGSDLDVPLIKLPTNTSDVQVYDLTLGSDGLFPQKGYMGFTWLDSPN-QT 1405  
Db 1370 dpngttfqpnlrsdqisllvlpdsskmqakivltdgiapqkgtgtltldpqlqn 1429  
QY 1406 GKLOARWTFDTYRRWYIPRDNHFYANSILGSONSMILVVKQGLINMLNNAFDDIAYNN 1465  
Db 1430 gtisalwkdsyrfqwyvprdnhfyansilgsmnvtvkqglndkmlnarfdevsynn 1489  
QY 1466 FWVSGVGTFLAQQGTPLSEEFYSYSRGTSVAIDAKPRQDPILGAAPSKIVGKTKAIAKMH 1525  
Db 1490 lwisglgtlmsqvgtpseefyysrgasvaldakpahdivgaafskmigtksikren 1549  
QY 1526 NYPHKGEYSYQASVYGGKFLPYLLNKQHGWAIPFIQGVVSYGHHKHDFTTLYPSIHER 1585  
Db 1550 nythkgeysyqasvygkpfhfvlnkktkslplllqgvisygykhdtvthyptirer 1609  
QY 1586 NKGWEDLGLADLRISMDLKEPSKSSKRITVYGELEYSIRIKOFTEDIDYPRFDDC 1645  
Db 1610 nggewedlglwltairvssvlrtpagqdkritvygeleysirkrqfteteypdryfnc 1669  
QY 1646 AYRNLSLPGCAVEGAIMNCNIMLKLALAYMPSIYRNPNVCKYRVLSNEAGQVIGCV 1705  
Db 1670 tyrnlaiplmglafegelesgndlimyrfsvaympsiyrnsptckyqvlssgegeilcgv 1729  
QY 1706 PRTSARAEYSTQLYLGPFWTLYGNTVIDGMVYTLQSMTSCGARMIF 1752  
Db 1730 ptrnsargeystqlvppilwtlygystieadahtlahmnmccgarmtf 1776

RESULT 4  
AAB13643  
ID AAB13643 standard; Protein; 948 AA.  
XX  
AC AAB13643;  
XX  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE C. trachomatis pmpC gene carboxy terminus protein.  
XX  
KW Chlamydial infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial.  
XX  
OS Chlamydia trachomatis.  
XX  
XX  
PN WO200034483-A2.  
XX  
PD 15-JUN-2000.  
XX  
XX  
PF 08-DEC-1999; 99WO-US29012.  
XX  
PR 08-DEC-1998; 98US-0208277.  
PR 08-APR-1999; 99US-0288594.  
PR 01-OCT-1999; 99US-0410568.

PR 22-OCT-1999; 99US-0426571.  
XX  
PA (CORI-) CORIXA CORP.  
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
XX  
XX WPI; 2000-431303/37.  
XX  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -  
XX  
PS Claim 2; Pages 216-218; 256pp; English.  
XX  
XX The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a protein isolated in the  
CC present invention.  
XX  
SQ Sequence 948 AA;  
Query Match 30.2%; Score 2703; DB 21; Length 948;  
Best Local Similarity 56.6%; Pred. No. 2.5e-129;  
Matches 543; Conservative 138; Mismatches 204; Indels 74; Gaps 16;  
QY 847 VTFSCQSGTCQFSGNAIDNNP--SQSSLNVOGGAIYAKTSLSIGSDAGTSYIFSCNSV 904  
Db 11 vkienfsgggifsgnkaidntteggssksnvlggavyaktlfnldgsgsrvtvfsgntv 70  
QY 905 STCKSQTTGCIAGAIYSPVTILNCIPATFNNNTASTATPKTSSSDGSSGNISDKDIGAI 964  
Db 71 ss--qsttgqavagayspvtlatpvyvfknsa-----tnnannatdtkrkdftgial 122  
QY 965 AGT-AITLGSVSRFSGNTADLGAAGTFLANANTPSATSGSONSTIKETILENGSFIIFERN 1023  
Db 123 gatsavsls99gahflenvadlgsaiglvpd-----tqn--tetcvlesgyfeyekn 171  
QY 1024 QANKGAIYSPSVSIRKNNITFNQNTSTHDGSAIYTKDATIESLGSLVFTGNNVTATQA 1083  
Db 172 kalkratiyapvvsikaytatfnqrslseegsaiyftkeasieslgslvftgnlvtpt-1 230  
QY 1084 SSATSGQNTNTA---NYGAAIFGDPGTTQSSQTDAL-LTLASSGNITFSNNLSLQNNQ 1138  
Db 231 sttetedpattsgdvtkygaalfiglasngsqtdnlplkliasggnicfrnneyrftss 290  
QY 1139 DTPASKFCSTIAGVYKLSLOAAKGTISFDCVHTSTKKTGSTQNVYETLDINKENS--- 1195  
Db 291 dtgtsrfsagiadvklmtqaakgtisfdairtktgtatqataydldinksedsetv 350  
QY 1196 -NPYTGTVFSSSELHENKSYIPQNALHNGTLVLEKTELHVVSFEQKESKLIMEPGAV 1254  
Db 351 nsaftgtlilfsselhenksyipqnvvlhsgsilvlpntelhlvisfeqegsalvntpgv 410  
QY 1255 LSNQNIANGALAINGLTIDLSSM---GTPQAGEIFSPPELRIVATTSASGG----- 1303  
Db 411 lsnqtvadgalvinmmtldlssveknglae-gniftppehrlidtttsgsggtptdses 469  
QY 1304 -----SGVSSSIP-----TNPKRISAAPVSGSAATTP---TMSE 1334  
Db 470 nqnsddtkceqnnndasngesangssspavaaahstrtrnfaaataatpttptattts 529  
QY 1335 NKVFLTGLTLIDPNGNFYQNPMGLSDLDVPLIKLPTNTSDVQVYDLTLGSDGLFPQKGYM 1394





Db 881 ahmncgarmif 892

RESULT 6  
AAAY37828  
ID AAAY37828 standard; Protein: 568 AA.  
XX AC  
XX AAAY37828;  
XX XX  
XX 07-OCT-1999 (first entry)  
XX XX  
DE Chlamydia trachomatis surface exposed protein.  
XX XX  
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;  
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;  
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.  
XX XX  
OS Chlamydia trachomatis.  
XX OS  
XX WO9928475-A2.  
PN PN  
PD 10-JUN-1999.  
XX PF  
XX 27-NOV-1998; 98WO-IB01939.  
XX PR  
XX 04-NOV-1998; 98US-0107077.  
PR PR  
XX 28-NOV-1997; 97FR-0015041.  
XX PR  
XX 17-DEC-1997; 97FR-0016034.  
XX XX  
PA (GEST ) GENSET.  
XX PA  
PI Griffais R;  
XX PI  
XX WPI; 1999-371125/31.  
DR DR  
XX Genome sequence of Chlamydia trachomatis  
PT PT  
XX Disclosure; Page 1399; 1755pp; English.  
XX XX  
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as  
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
CC can also be used to control growth of the microorganism. Chlamydia  
CC trachomatis is responsible for a large number of diseases, e.g. eye  
CC diseases such as conventional trachoma, nonendemic trachoma,  
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
CC nongonococcal urethritis, epidymitis, cervicitis, salpingitis,  
CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;  
CC and venereal lymphogranulomatosis. The polypeptides of the invention  
CC may be of use in treating these diseases.  
XX XX  
SQ Sequence 568 AA;

Query Match 28.7%; Score 2568.5; DB 20; Length 568;  
Best Local Similarity 93.4%; Pred. No. 8e-123;  
Matches 512; Conservative 14; Mismatches 21; Indels 1; Gaps 1;

QY 15 PSVSGCFPEPKELNFSRVTSSSTTTTETIGEGAIEYIVSGNASFTKFTNIPDTTTP 74  
| |  
| |  
Dd 14 pqfkgfasqnlxxlnfsrvgtssstttftvgeagaylvssnsasfckftnptntttp 73

QY 75 TNSNSSSSGGETASVEDSDSTTTTTTPDKGGAFYNNAHSGVLSPWTRSGTGSLTLSEIK 134  
| |  
| |  
Dd 74 tnsnsnsretasvksdsdttttkpdkggafynahsgvlsmtrsgtgsltlseik 133

QY 135 MTGEGGAIFSQGELLFDTLSLTQNQLSQLSGAIFGGSTISLSGITTKATFSCNSAEVP 194  
| |  
| |  
Dd 134 mtgeggaifsqgellfdlglgtiqnnlsqsgaifggstislsgitkatfscnsaevp 193

QY 195 APVKKPTPEPQAQTASETSGSSSGNDPSVSSPPSARPEAAPANTLOSHFICATATPAAOQT 254  
| |  
| |



CC It is an example of C. pneumoniae antigenic polypeptides of the  
CC invention (see AA95543-55) that are encoded by open reading frames  
CC (see AAA50030-42) identified in the C. pneumoniae genome. The  
CC polypeptides are useful in the diagnosis, treatment and prevention  
CC of Chlamydia infection. They can be prepared by recombinant  
CC methods using transformed unicellular host cells. The  
CC polypeptides, or immunogenic fragments of them, or fusion proteins  
CC that include an additional polypeptide such as a heterologous  
CC signal peptide or a polypeptide having adjuvant activity, are  
CC utilised as vaccines to protect against Chlamydia infection.  
CC The polypeptides are also used in diagnostic kits and in methods  
CC of detecting Chlamydia infection.  
XX  
XX  
SQ Sequence 469 AA;

Query Match 14.2%; Score 1272.5; DB 21; Length 469;  
Best Local Similarity 47.3%; Pred. No. 4.3e-57;  
Matches 239; Conservative 89; Mismatches 140; Indels 37; Gaps 3;  
QY 1249 MPEGAVLSNQNTANGALAINGLTIDLSMGTQAGEIFSPPELRIVATVTSASGGSGVSS 1308  
Db 1 mppgsvlnshsksaegiainnvildseivptkdnatvapptklvrtnad----- 52  
QY 1309 SIPTNPKRISAAVPSGSAATPTMSKNVFLGDLTLPDNGNFYQNPMLGSDLDVPLIK 1368  
Db 53 -----skdkiditgtvlldpngnlyqnsylgedrditln 88  
QY 1369 LPTNTSD-VQVYDLTSLGDLFPQKQVMGWTLDSPQTKLOARWTFDNRWVYIPRON 1427  
Db 89 idnsasgavtatnvtlqgnlqakkgylgtwnldpnssgskllkwtfdkylrwpypdn 148  
QY 1428 HFYANSILGSQNSMIVKQGLNNMLNNAFDDIAYNNFWVGVTFLAQOQTPLSEFS 1487  
Db 149 hfynsiwgaqslvtvngilgnlnnarfedpafnnfwsaigsflrkevsrnsdft 208  
QY 1488 YVSRGTSVAIDAKPRODFILGAFAFSKIVGKTKRAIKKMHNFHKGSYSYQOASVYGGKFLY 1547  
Db 209 yhgrytaavdakprqefilgaafsqvfghaeseyhldnykkgshstqaslyagnify 268  
QY 1548 FULLNQHQWALPFLQGVVSYGHIKHTDTTLVPSIHERNKGWEDLGWLADLRISMDLKE 1607  
Db 269 fpaistr-----piligvfatygmqndtttlypsaeknmanwdsawldfrfsvdlke 324  
QY 1608 PSKDSKRITVCELEYSSIROKQTEIDYDPRHFDCCAYRNLSLPVGCAGEAIVMNCNI 1667  
Db 325 paphstarltfyeaeayrirqekftelydprfsacsygnlaipgtgfsvdgalawrei 384  
QY 1668 LMVNLALAYMPSIYRNPNVCKYRLSSNEAGOVICGVPTRTSARAEYSTQLYLGPFWTL 1727  
Db 385 ilynkvsaaylpvlirnpkayevlstkekgnvvnlptrnaaraevssqilylgsywtl 444  
QY 1728 YGNVYTDVGMVTLTSMQTSAGRMIF 1752  
Db 445 ygcytidasmntvlqmganggrifrv 469

RESULT 9  
AAV37823  
ID AAV37823 standard; Protein; 915 AA.  
XX  
AC AAV37823;  
XX  
DT 07-OCT-1999 (first entry)  
XX  
DE Chlamydia trachomatis cellular envelope protein.  
XX  
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;  
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.  
XX  
OS Chlamydia trachomatis.

XX WO9928475-A2.  
PN 10-JUN-1999.  
XX  
PD 27-NOV-1998; 98WO-IB01939.  
XX  
PF 04-NOV-1998; 98US-0107077.  
XX  
PR 28-NOV-1997; 97FR-0015041.  
PR 17-DEC-1997; 97FR-0016034.  
XX  
PA (GEST ) GENSET.  
XX  
XX Griffais R;  
PI  
XX WPI; 1999-371125/31.  
DR  
XX  
XX Genome sequence of Chlamydia trachomatis  
PT  
XX Disclosure; Page 1391-1393; 1755pp; English.  
PS  
XX AAV36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
CC can also be used to control growth of the microorganism. Chlamydia  
CC trachomatis is responsible for a large number of diseases, e.g. eye  
CC diseases such as conventional trachoma, nonendemic trachoma,  
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
CC perihpatitis, bartholinitis; pneumopathy in breast feeding infants;  
CC and venereal lymphogranulomatosis. The polypeptides of the invention  
CC may be of use in treating these diseases.  
XX  
SQ Sequence 915 AA;  
Query Match 11.2%; Score 1002; DB 20; Length 915;  
Best Local Similarity 27.3%; Pred. No. 6.3e-43;  
Matches 331; Conservative 153; Mismatches 323; Indels 404; Gaps 34;  
QY 1 MWLSTAYFAAVALPSVSGFCPEPK-----ELNFSRVETSSSTTTTETIGEAGAEIVSG 56  
Db 2 mkflstatafaavalsvteassiqdqikntdcnvskvgystsqafdmldadnteayraad 61  
QY 57 NASFTKFTNIPDTTPTPTNSNSSSGTASVSDSTTTTPDPKGGGAFNAHSGVL 116  
Db 62 svsfydf-----stsglprkhlssseasptt-----egv- 92  
QY 117 SPMTRSGTEGSLTSLSEIKMTGEGGAIFSQGELLFTDLTSLTIQNNLSQLSGGAIFGGSTI 176  
Db 93 -----sssggentenagds----- 92  
QY 177 SLUGITKATFSCNSAEVPAPVKKPTPEKQATASSETSGSSSSSGNDSDVSPSSRAEPAA 236  
Db 93 -----sssggentenagds----- 107  
QY 237 NLQSHFICATATPAAOTDTETSPSHKPGGGAIVAKGDLTIADSO----- 282  
Db 108 -----apssgetdkkt-----eeeldnggiyirekltisesqdslnpsielhdns 154  
QY 283 -----EVLFSINKATKGGGAIFAEXKDVSEFNITSLKVQTNGAEEKGGAIVAKGDLISQS 336  
Db 155 fffgegevfidhrvalknggalygekvfvfenlksllvevnisxekgsvayakervslen 214  
QY 337 SKQSLFNSNYKQGGGALYVEGINFQDLEETRIKYNKAGCTFETKKI----- 383  
Db 215 vteatfssnggggggilyseqdmllisdcnvfhfgnaagatavkqcldeemilvitacv 274  
QY 384 -----TLPSLKAQAAGNADAWASSPQSGGATTVSDGSSGSDSDSTST/VPY 434  
Db 275 dalsedtltpeteqtksnqndg-----ssetkdtqvsespesptpdd-----v 321  
QY 435 TAKGGGLYTDKNLSITNITGIETIANNKATDVGGGAYVKGTTLTCENSHRLQLKNSSDKQ 494

Db 322 lqkggyitekstlgtitgtdfvsntatdsagvftkenlsctntnslgflknsag-- 379  
Qy 495 GGGIYGEDNITLSNLTGKTLQENTAKEEGGLFKFGDKAL/TWGTGLDSFCLINNTSEKH 554  
Db 380 -----qg 381  
Qy 555 GGGAFVTRKISQT-YTSDVETPIGTPVHGTVITGNKSTGGNGGVCTKRLALSNLQSI 613  
Db 382 gggayvtqmtsvnttsesittp---plvgevifsen-takghggictnklslsnlktv 437  
Qy 614 SLSGNSAAENGGAHTCPDSEPTTADTAQAPAAASATTPKSPAPVSTALSPSSSTVSSSL 673  
Db 438 tltnksakesgaiftldaspttdpes-----stpsssspast 477  
Qy 674 TLLAAS-----SQASPA-----TSNKETODPNADTDLIDYVVDITISKNT 714  
Db 478 xevvasaklnrfastaepaapslteasdqtdqtetsdtnsdldvsnlinvainqnt 537  
Qy 715 -AKKGGGIYAKKAKMSRIDQLNISSENSATEIGGGICCKESLELDALVSLSVTENLVGKEG 773  
Db 538 sakkgaigykaklsrinnlielsnsgqvgggllcltesvefdaigslshynsaakeg 597  
Qy 774 GGLHAKTVNINLKSFGFSNN--KANSSSTGVATTASAPAAAAAASLQAAAAAAPS----- 827  
Db 598 gvihskvtlslnktftfadntvkaivestpeapeeippvege---estatempnsnte 654  
Qy 828 -SPATPTYSGVVG-----GAIYGEKYTFSCQSGTCQFSGNQADNPNPSQSLNVQGA 879  
Db 655 gssantnlegsgqtdadtgtgvvnnesqtdsd-tgnae-sgeqlqdstsneentipnss 712  
Qy 880 IVAKTSLSIGSSDAGTSYIFGNSVSTGKSQTTGQIAGGAIYSPRTVTLNCPA---TFSN 936  
Db 713 lqsnentdesshteeitdesvsskskgsstpqdgaass-----gapsgdgsisan 767  
Qy 937 TASIATPKTSSDGSNGSINKDTIGGATAGTATILSGVSRFSNGTADIGAAIGTLANNT 996  
Db 768 aclaksyaastdspsvns-----sgsdvtassndpdsessgd-----sagds 810  
Qy 997 PSATSGSNSTIEKTLLENGSFIFERNQANKRGAIYSPSVSTIKGNITFNQNTSPHDGSA 1056  
Db 811 eptpepgsttetptlllg-----galygetvkl-----enfs----- 844  
Qy 1057 IVFTKDATIESLGSLVLTGNNVTATQASSATSGQNTNTANYGAAIFG-----DPGT-TQ 1109  
Db 845 -----gggifsngka-----idntteggssksnvlvggavyaktllfnldsgtide 888  
Qy 1110 SSQTDAILTLL 1120  
Db 889 lpspgilsll 899  
RESULT 10  
AAB13644  
ID AAB13644 standard; Protein; 821 AA.  
XX AC AAB13644;  
XX AC AAB13644;  
DT 02-FEB-2001 (first entry)  
XX DE C. trachomatis pmpC gene amino terminus minus signal sequence protein.  
XX KW Chlamydial infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial.  
XX OS Chlamydia trachomatis.  
XX PN WO200034483-A2.  
XX PD 15-JUN-2000.  
XX.

PF 08-DEC-1999; 99WO-US29012.  
XX 08-DEC-1998; 98US-0208277.  
PR 08-APR-1999; 99US-0288594.  
PR 01-OCT-1999; 99US-0410568.  
PR 22-OCT-1999; 99US-0426571.  
XX (CORI-) CORIXA CORP.  
PA Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
PI WPI; 2000-431303/37.  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
XX comprises immunogenic portion of Chlamydia antigen, which comprises  
XX amino acid sequence encoded by polynucleotide sequence -  
XX Claim 2; Pages 219-220; 256pp; English.  
XX The present invention relates to new nucleic acid sequences and the  
XX proteins encoded by the nucleic acid sequences. The encoded proteins  
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded  
XX proteins are useful for the serodiagnosis and treatment of Chlamydia  
XX infection. Chlamidia are intracellular bacterial pathogens that are  
XX responsible for a wide variety of human infections. C. trachomatis  
XX infection is one of the most common sexually transmitted diseases and can  
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
XX and infertility. Trachoma due to ocular infection with C. trachomatis is  
XX the leading cause of preventable blindness worldwide. C. pneumoniae is a  
XX major cause of acute respiratory tract infections in humans and is also  
XX thought to play a role in the pathogenesis of atherosclerosis and  
XX coronary heart disease. The present sequence is a protein isolated in the  
XX present invention.  
XX Sequence 821 AA;  
SQ  
Query Match 10.6%; Score 943.5; DB 21; Length 821;  
Best Local Similarity 31.1%; Pred. No. 5e-40;  
Matches 277; Conservative 137; Mismatches 283; Indels 195; Gaps 27;  
Qy 181 ITKATFSCNSAEVPAPVKKTEPKAQT-----SETSG-----SSSSSNDSSVSSPS 228  
Db 23 vskgygstsqftdmldadnteyraadsvsfydfstssglprkhlssseaspttegvss 82  
Qy 229 SRAPPAANLQSHFICATATPAQOTDTETPSHKGSGGAIYAKGDLTIADSQ----- 282  
Db 83 sssgentensqd-----sapssgetdkkt---eeeldnggilyarekltisesqdslnp 134  
Qy 283 -----EVLFSINKADKGGAIFAEKDVSPENITSLKVQTNGAEEKGAIYA 328  
Db 135 sielhndnsffgegevidhrvalknnggalygekevfvfenikslivevnisvekggsyva 194  
Qy 329 KGLDLSIQSSKOSLFSNSYKOGGALYVEGGINFODLEIRIKYNKAGCTFTFKTI----- 383  
Db 195 kervslenvteatfssnggegggyiseqdmllisdcnvfhfgnaagatavakqcldeem 254  
Qy 384 -----TLPSLKAQASAGNADAWASSFPQSCSGATTYVSDSCDSSGSDS 426  
Db 255 ivlltecvdsisedtldstpeteqtknsngdqg-----ssetkdtqvsespstpsdd 308  
Qy 427 DTSETVPYTAKGGLYTDKNLSITNITGIIETANNKADVDGGVAYVKGTLCENSHRLQF 486  
Db 309 -----vlgkggyitekstlgtitgtdfvsniatdsagvftkenlsctnslsqf 361  
Qy 487 LKNSSDKOGGGIYGEDNITLSNLTGKTLFQENTAKEEGGLFIKGTDKAL/TWGTGLDSFCL 546  
Db 362 lknsgag-----  
Qy 547 INNTSEKHGGGAFVTRKISQT-YTSDVETPIGTPVHGTVITGNKSTGGNGGVCTKRL 605  
Db 368 -----qhggyayvtqmtsvnttsesittp---plvgevifsen-takghggictnkl 417

QY 606 ALSNLSISISGNSAENGGGAHTCPDPTADTARQ---PAAASAANTSTPK---SAPVST 660  
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 476  
s|an|k|t|v|l|t|k|n|s|a|k|e|s|g|a|i|f|d|l|a|s|p|t|d|t|p|e|s|t|p|s|s|p|a|s|t|p|e|v|v|a|s|a|k|i|n|r  
QY 661 ALSTPSSSTVSSLTLLAASSQASPATSNKETODPNADTDLIDYVVDTTISKNT-AKKG 719  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 531  
f|f|a|s|t|a|e|a|p|a|s|l|t|-----e|a|e|s|d|q|d|t|e|t|s|d|n|s|d|d|v|s|i|e|i|n|l|n|v|a|i|n|g|n|t|s|a|k|k|g  
QY 720 GIYAKAKMSRIDQLNISNSATEIGGICCKESLELDALVLSVTENLVKRGGLHAK 779  
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 591  
a|y|g|k|k|a|k|i|s|i|n|n|l|e|l|s|n|g|s|g|d|v|g|g|i|c|l|t|e|s|v|e|f|a|i|g|s|i|l|l|s|y|n|s|a|a|k|e|g|v|i|h|s|k  
QY 780 TVNINSLKSGFSPNN--KANSSSTCVATTAAPAAAAAASLQAAAAAPS-----SPATP 832  
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 648  
t|v|l|s|n|l|k|t|f|a|d|n|t|v|k|i|a|e|t|p|e|a|i|p|p|v|e|g|e|-----e|s|t|a|t|e|i|n|p|n|s|t|e|i|g|s|a|n|t  
QY 833 TVSGVVG-----GAIYGEKVTFSOCSTCFSGNQAIIDNPNQSSSLNVQGGAIYAKTS 885  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 706  
n|l|e|g|s|g|d|t|a|d|t|g|t|v|v|n|n|e|s|q|d|s|d|t|g|n|a|e|s|g|e|l|q|d|s|t|q|s|n|e|i|t|p|n|s|i|d|q|s|n|e  
QY 886 LSTGSSDAGTSYIFSGNSVSTGKSTQTGTAGGAIYS-----PTVTLNC----- 929  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 766  
n|d|e|s|s|d|s|t|e|i|l|d|e|s|v|s|s|k|s|g|s|t|p|q|d|g|a|a|s|g|a|s|g|q|s|i|s|a|n|a|c|i|a|k|s|y|a|a|s|t  
QY 930 -PATFSNNTASTATPKTSBED-GSSGNSIKDTIG-----GAITAGTAITLSG 973  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 818  
d|s|p|v|n|s|g|s|d|v|t|a|s|s|d|n|p|d|s|s|g|s|d|s|g|d|s|e|g|t|e|p|e|a|g|s|t|e|t|p|i|l|g|g  
RESULT 11  
AAV35617  
ID AAV35617 standard; Protein; 1617 AA.  
XX AAV35617;  
DT 13-SEP-1999 (first entry)  
XX Chlamydia pneumoniae surface exposed polypeptide.  
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX Chlamydia pneumoniae.  
OS W09927105-A2.  
PN 03-JUN-1999.  
XX 20-NOV-1998; 98WO-IB01890.  
XX 04-NOV-1998; 98US-0107078.  
PR 21-NOV-1997; 97FR-0014673.  
XX (GEST ) GENSET.  
PA Griffais R;  
XX WPI; 1999-357842/30.  
XX Genome sequence of Chlamydia pneumoniae  
PS Page 1344-1347; Disclosure; 1912pp; English.  
XX AAV34584-V35879 represent the proteins encoded by all the open reading  
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAV34584-V35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,

CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.  
XX  
SQ Sequence 1617 AA;  
Query Match 8.4%; Score 748.5; DB 20; Length 1617;  
Best Local Similarity 22.3%; Pred. No. 1.1e-29;  
Matches 406; Conservative 267; Mismatches 685; Indels 459; Gaps 79;  
QY 165 LSGGAIFGSGTISLS---GITKATFSCNSA---EVPAPVKKPTPE--KAQTASE----- 210  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 90  
l|s|a|i|a|f|e|a|h|s|h|s|e|i|d|i|g|v|f|n|k|f|e|h|s|a|h|v|e|e|a|q|t|s|v|l|k|s|d|p|v|n|s|q|k|e|k|v|y|t  
QY 211 ---TSGSSSS---GND-----SVSSPSSSFAEP 233  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 147  
q|v|p|l|t|g|s|g|e|i|d|a|n|x|l|h|f|q|h|l|f|e|t|t|v|f|i|d|q|k|l|v|s|d|l|t|r|n|f|s|q|t|q|---e|p  
QY 234 AAANLQSHFICATPATPAQTDTETSTPSHKPG---SGGAIYAKGDLTIAADQEVLFISIN 289  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 206  
d|t|s|n|a|v|s|e|i|l|s|s|d|t|k|e|r|k|d|l|e|d|p|s|k|k|g|i|k|e|v|s|d|i|p|k|s|p|e|t|a|v|a|i|s|e|i|d|e|i|s|e|i|n  
QY 290 KATKD---GGAIF-----AEKDVSPENI-----TSLKVQTNAGEEKGAIIYAK 329  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 266  
i|s|a|r|d|p|l|q|l|a|f|f|y|k|n|t|s|s|i|e|k|s|s|f|q|i|i|f|s|g|a|n|s|i|g|i|f|e|i|n|k|p|k|s|g|a|a|v|s|d  
QY 330 GDLSIOSQSQSL-FNSNYSKQGGALYVEGGINF-----QDLEEIFI-- 370  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 322  
r|d|i|v|f|e|i|n|v|k|i|s|i|c|e|i|s|e|i|d|g|s|a|---a|g|i|v|i|v|t|h|g|d|v|t|i|d|c|a|t|g|i|d|l|e|i|r|i|v|k  
QY 371 KYNKAGTFETK-----ITPLSLKAQASAGNADAWASSPQSGGATTVSDGSSSG 423  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 382  
d|f|s|r|g|g|a|v|f|a|r|i|n|h|e|v|q|i|n|l|a|g|i|l|s|v|v|g|i|k|a|i|v|v|e|i|k|s|a|e|i|n|g|i|n|i|g|i|n|s|---g|i|a|i|t|f  
QY 424 SDSDTSETVPVTAAGGLYTDKNLSITNITGIIETIANNKA-----TD-VGGGAY- 471  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 440  
n|t|a|l|w|k|e|i|n|q|i|s|---g|i|a|i|s|a|i|d|i|q|i|n|c|i|a|i|e|i|f|g|i|n|q|i|s|i|a|i|g|i|h|i|g|i|t|d|i|f|v|i|g|g|i|a|l  
QY 472 VKGTLCENSHRQFLKNSDKOGGIGYEDNITLNLTKTLFOENTAKEEGGFLIKG 531  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 497  
a|q|g|t|i|l|r|i|n|n|a|v|v|q|v|k|t|s|k|h|g|i|a|i|l|a|---g|i|v|d|i|n|e|i|s|e|i|v|a|f|k|i|n|t|a|i|l|t|g|i|a|l|---s|a  
QY 532 TDKALMTGLDSFCLINNTSEKHG---GAFVTKISQTYTSDVETIPGITPVHGETVI 587  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 554  
n|d|k|i|i|i|a|i|n|f|e|i|l|f|e|i|q|e|v|e|r|i|n|h|g|i|a|i|y|c|r|i|n|p|k|l|e|q|x|i|g|i|n|i|i|g|i|n|s|---g|i|a|i|t|f  
QY 588 TGNKST-----GGNGGVCTKRLAL-SNLQISISGNSAEN-----GGAHT 629  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 614  
l|i|k|k|a|s|v|l|e|v|t|q|a|e|d|y|i|g|i|a|i|l|g|h|i|v|i|l|d|i|s|i|n|i|q|i|f|i|g|i|n|i|g|i|n|f|i|g|i|v|y|i|g|i|a|i|l  
QY 630 CPDSFPTADTA-----EQPAAASAATSTPKSAPVSTALSTPSSSTVSSLTLLAASSQ 681  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 673  
s|t|d|r|v|i|s|i|n|i|n|s|g|d|v|f|i|k|i|n|g|k|q|c|i|a|q|y|i|a|p|e|i|t|a|p|i|e|i|s|i|d|e|i|s|---t|i|k|d|e|i|s|i|n|a|c|h|i|d|h  
QY 682 ASPATSNKETQDPNADTDLIDYVVDTTISKNTAKKGGIYAKKAKMSRIDQLNISNSA 741  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 717  
y|i|p|k|t|v|e|i|e|---v|i|p|i|s|i|l|e|i|h|p|v|---v|i|s|t|i|d|i|r|g|i|g|i|a|i|l|a|q|---h|i|f|i|d|e|i|t|g  
QY 742 TEIGGGICCKESLELDALVLSVTENLVKEGGLHAKTV-NISNLKSGFSPNNKAN-S 799  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 754  
n|i|r|f|i|s|g|i|n|l|---g|i|g|i|e|i|s|t|v|i|g|i|a|i|v|i|g|i|g|i|a|i|l|l|t|s|i|n|e|i|v|i|c  
QY 800 SSTGVATTASAPAAAAAASLQAAAAAPSSPATPTYSGVVGAIYGEKVTFSOCSTCFES 859  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 784  
s|i|q|i|n|v|i|f|i|d|i|v|i|n|g|i|c|i|d|i|s|---g|i|a|i|l|a|k|k|v|-----d|i|s  
QY 860 GNOQAI--NNPQSSSLNVQGGAIYAKTSLISGSSDAGTSYIFSGNSVSTGKSTQITQIAG 917  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 830  
a|h|i|s|i|v|e|i|f|i|v|i|g|i|g|i|f|---g|i|g|i|v|i|c|i|n|e|---s|i|n|i|t|i|d|---n|i|g|i|s|i|v|i|f|i|s|i|k|i|n|t|---i|i|g|i|g  
QY 918 GAIYSTVTILNCPATPSNNTASTATPKT---SSEDESSGNSIKDTIGGAIAGTAITL--- 971  
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 880  
a|g|i|v|i|a|p|i|q|i|g|i|v|---t|i|c|i|g|i|n|i|g|i|a|i|f|i|n|f|i|g|i|s|i|e|i|r|i|g|i|g|---g|i|i|i|a|i|n|i|s|i|v|i|i|q|i|n



```
Db 306 dkgtlsgnqemfsnttttngangaiyatkctldntltfdqnta-----tagc 356
Qy 659 STALSTPSSSTVSSLTLLAASSOASPATSNKETQDPNADTDLIDYVDTTTSKNTAKKG 718
Db 357 ggaityt-----etedsf-----lkgstglvtfstntakg 386
Qy 719 GGIYAKKAKMSRIDQLNISENATEIGGICCKESLELDALVSLVTENLVNKEGGGLHA 778
Db 387 galyk-----ens-----slgtntnll-----404
Qy 779 KTVNISNLKGFSGFNKANSSTGVATTASAPAAAAAASLOAAAAAPSSPATPTYSGVV 838
Db 405 -----fsgnktgpnssanegcggallsflesasvt-----438
Qy 839 GGAIYGEKVTFQSCGTCGFSQNGAIYDNNPSSQSLNVOGAIYAKTSLSIGSSDAGTSYI 898
Db 439 -----kglwiednevnslgnta-----tvsggaiyatkcalthgn-----ttlt 478
Qy 899 FSGNSVSTGKSQTTGOIAGAIYSPVTVINC-----PATFSNNTASIAPTKSSSEDCSSG 953
Db 479 fdgn-----taetaggaiytetedftltgstgtvtfstnta-----ktagalhthg 524
Qy 954 NSIKDTIGGAIAGTATLTLGVSFRFSGNTADLGAAIGTLANANTPSATSGQNSI-----1007
Db 525 n-----tsfthkalfvsnsa-----tatatttdqecgcallcnise 564
Qy 1008 ----TEKITL-ENGSPIFERNQANKR-GAIYSPSVSIKGN-NITFNQNTSTHDGSAIYFT 1060
Db 565 sdiatksltitenesisfinntakrsgglyapkvvisgesinfngntaetsggaiy-s 623
Qy 1061 KMTIESLGSVLTGNNVTATONASSATSGONTNANYGAIFGDPGTTTOSSQTDAILTL 1120
Db 624 kalsitangpvsftnn-----sgkg-----gaiyadsge-----lsle 658
Qy 1121 ASSGNITFSNLSLONNQDTPASKFCSIAGYVKLSLQAAGKTIISFDCVHSTKTKGST 1180
Db 659 aldgdtlftfsgn-----rategstpnshlgagakitkilaapghgtiifydptmeapasgt 716
Qy 1181 QNVYETLDIN-----KEEN-----SNPYTGTVIFSS-ELHENKSYIP 1216
Db 717 ---ieelvnpvkaivppqkngpiasvvpvapaupntgtliffssgkipsqdasip 773
Qy 1217 QNAI-----LHNGTILVLKEKTELHVVSFEQESKLIMEPAVL--SNQNIANGALA 1266
Db 774 antttlndqkinlaggnvvlkegatlqvysftqppdstvmdagttletttntntdgsid 833
Qy 1267 INGLTIDLSSMGTPOAGEIFSPPELRIVATTSSAGSGSVSSSIPTNPKRISAAVPSGSA 1326
Db 834 knlsvnlaldgkr-----mitiavnstsgglkis-----864
Qy 1327 ATPTMSENKVFILTDLTLDPNGFYQNPMLGSDLDVPLIKLPTNTSDVQVYDLR--LS 1384
Db 865 -----gdlkfhnnegsfydnpglkanlnlpfldlsstsgtgnlddfnpps 910
Qy 1385 GDLFPQKGYWGTWTLDSNPTQC-----KLOARWTFDYR-----RWVYIPRONHFTYANSI 1434
Db 911 smaapdygvqgswtl--vpkvagagkvtiavaeqalgytpkpelratlvp-----nsl 961
Qy 1435 LGSQNSMIVVVKGLINMLNNAFFDDIAYNNFWVSGVGFFLAQCGTPLSEEFYSYSRGTS 1494
Db 962 wnayvnihsiqeia-----tamsdapsbgwiigisgnafhqdkkenagflisrgyl 1016
Qy 1495 VAIDAKPRODFILGAFAFSKIVGKTK--AIKKMHNYPFKGS--EYSYQASVYVG--KFL 1546
Db 1017 vggsmtpdqeytfavafslfkskdyvvsdiksyyagalscaqssyviplhsrlrrhl 1076
Qy 1547 YFLLNKQHGVALPFLIQGVVSVGHIKHDTTTLTPSPSIHERNKGDWE 1591
Db 1077 skvlpelpg-etplvlhgvsvygrnhhmttklan-ntgkgsdwd 1119
```

RESULT 13  
AAB13635

```
ID AAB13635 standard; Protein; 1530 AA.
XX
AC AAB13635;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis pmpD gene protein.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US29012.
XX
PR 08-DEC-1998; 98US-0208277.
PR 08-APR-1999; 99US-0288594.
PR 01-OCT-1999; 99US-0410568.
PR 22-OCT-1999; 99US-0426571.
XX
(CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
WPI; 2000-431303/37.
XX
DR Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises
PT amino acid sequence encoded by polynucleotide sequence .
XX
PS Claim 2; Pages 186-190; 256pp; English.
XX
CC The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamidia are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention.
XX
SQ Sequence 1530 AA;
XX
Query Match 7.7%; Score 685.5; DB 21; Length 1530;
Best Local Similarity 21.7%; Pred. No. 1.6e-26;
Matches 417; Conservative 251; Mismatches 670; Indels 583; Gaps 82;
QY 8 AVPAALVPSVSGRC-----FPEPKELNFSRVETSSSTFT 42
Db 17 svvaailasvsglascvdlhaggvgvnelvyvgpqavllldqirdlrfvsgksqaeqlyr 76
QY 43 ETIGEAGAEIVSGNASFTKFTNIPRTDITPTI-----NSNSSSSSGETAS-VSESDS 95
Db 77 llvgdpsfqekdadtlpgkveqstlfsvtnpvnvfgvqdgqdvsgglicsfstsnlds 136
QY 96 TTTTPDPKGGAFYNAHSAHSGVLSFPMRTSGSLTLEIKMTGEGGAIFSGOGLLFTDLTS 155
Db 137 -----prdgesflg-----iafvgdsskag-iltldvksalgaalystedlif----- 179
QY 156 LTIONNLSOLSGGAIFGG-STISLSGITKA-----TFSCNSAEVPAPVKPTEPKAQTASE 210
```









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: October 25, 2001, 08:39:06 ; Search time 48.31 Seconds  
(without alignments)  
4798.145 Million cell updates/sec

Title: US-09-454-684A-180  
Perfect score: 8942  
Sequence: 1 MKWLSATAVFAAFLPSVSGF.....IDVGMVTLQMTSCGARMIF 1752

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_16.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_unclassified.\*  
13: sp\_vertebrate.\*  
14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	8827.5	98.7	1751	2	O84418
2	6488.5	72.6	1672	2	Q9PJY2
3	3550	39.7	1770	2	O84419
4	3205	35.8	1460	2	Q9PJY1
5	3112.5	34.8	1723	2	Q92812
6	3111.5	34.8	1723	2	Q9RB59
7	3111.5	34.8	1732	2	Q9K2C1
8	932.5	10.4	1407	2	Q92899
9	823.5	9.2	1276	2	Q9JRW2
10	743.5	8.3	1609	2	Q9RB58
11	743.5	8.3	1609	2	Q926U5
12	674	7.5	1531	2	O84818
13	665.5	7.4	930	2	Q9RB66
14	661.5	7.4	930	2	Q92393
15	655.5	7.3	914	2	O86163
16	655.5	7.3	928	2	Q9RB65
17	622.5	7.0	1520	2	Q9PLB0
18	579.5	6.5	926	2	P71135
19	577	6.5	2232	5	P91365

20	567	6.3	847	2	P71132
21	563.5	6.3	846	2	P71133
22	559	6.3	839	2	P77792
23	555.5	6.2	936	2	O92898
24	553.5	6.2	936	2	O9JSA2
25	530.5	5.9	922	2	Q929G5
26	529.5	5.9	922	2	Q924H9
27	528.5	5.9	922	2	O9K1Y9
28	517.5	5.8	949	2	O9K299
29	516.5	5.8	928	2	O86164
30	516	5.8	975	2	O84417
31	509	5.7	772	2	O9RB71
32	506	5.7	841	2	O923A1
33	504.5	5.6	2586	5	O9VTK8
34	504	5.6	1013	2	O84879
35	495	5.5	928	2	O92398
36	495	5.5	928	2	O92398
37	470	5.3	986	2	O9PJY3
38	470	5.3	973	2	Q92896
39	467.5	5.2	3381	2	O9K2A1
40	466	5.2	3381	2	O9KX33
41	461.5	5.2	978	2	O84880
42	459.5	5.1	2178	2	O9KWR3
43	456.5	5.1	978	2	Q92895
44	448.5	5.0	987	2	O9PL45
45	446.5	5.0	1275	5	O76602

#### ALIGNMENTS

RESULT 1  
O84418 PRELIMINARY; PRT; 1751 AA.  
AC O84418;  
DT 01-NOV-1998 (TremBLrel. 08, Created)  
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)  
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)  
DE PUTATIVE OUTER MEMBRANE PROTEIN B.  
OS PMPB.  
GN Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
ON NCBI\_TaxID=813;  
RX STRAIN=D/UV-3/CX;  
RX MEDLINE=99000809; PubMed=9784136;  
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;  
RA "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis."  
RL Chlamydia trachomatis."  
DR EMBL; AE001314; AAC68010.1; -  
DR InterPro; IPR003357; -  
DR Pfam; PF02385; OMP; 1.  
DR Pfam; PF02415; DUF145; 1.  
SQ SEQUENCE 1751 AA; 183011 MW; C6BE3F44DA7618B3 CRC64;

Query Match 98.7%; Score 8827.5; DB 2; Length 1751;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1732; Conservative 8; Mismatches 10; Indels 3; Gaps 2;

QY 1 MKWLSATAVFAAFLPSVSGFCPEPKELNFSRVETSSSTFTTETIGEAAGAEYIVSGNASF 60  
Db 1 MKWLSATAVFAAFLPSVSGFCPEPKELNFSRVETSSSTFTTETIGEAAGAEYIVSGNASF 60  
QY 61 TKFTNPTTDTTPTNSNSSSSGSETASVSDSDSTTTTDPKGGGAFYNAHSGVLSFWT 120  
Db 61 TKFTNPTTDTTPTNSNSSSSGSETASVSDSDSTTTTDPKGGGAFYNAHSGVLSFWT 120

```
Qy 121 RSGTEGSLTLSEIKWTEGGAIFSGQELLFTDLTSLATQNNLSQSGGAIFGGSTISLSG 180
Db 121 RSGTEGSLTLSEIKITGEGGAIFSGQELLFTDLTGLTQNNLSQSGGAIFGEGSTISLSG 180
Qy 181 ITKATFSCNSAEVPAPVKPTEPKAQTASETSGSSSSGNDVSPSSSRAEPAANLQS 240
Db 181 ITKATFSSNSAEVPAPVKPTEPKAQTASETSGSSSSGNDVSPSSSRAEPAANLQS 240
Qy 241 HFICATAPAAQTDETSTPSHKPGSGGAIYAKGDLTADSOEVLFSINKATKGGALFA 300
Db 241 HFICATAPAAQTDETSTPSHKPGSGGAIYAKGDLTADSOEVLFSINKATKGGALFA 300
Qy 301 EKDVSFENITSLKVQTNGAEKGGAIYAKGDLISQSSKQSLFNSYKQGGGALYVEGGI 360
Db 301 EKDVSFENITSLKVQTNGAEKGGAIYAKGDLISQSSKQSLFNSYKQGGGALYVEGGI 360
Qy 361 NFQDLEETRIKYNKAGTETTKITLPSLKAQASAGNADAWASSPQSGGATTVSDSDS 420
Db 361 NFQDLEETRIKYNKAGTETTKITLPSLKAQASAGNADAWASSPQSGGATTVSDSDS 420
Qy 421 SSGSDSDTSETVPVTAAGGGLYTDKNLSITNITGIIETANNKATDVGGGAYVKGTLTCEN 480
Db 421 SSGSDSDTSETVPVTAAGGGLYTDKNLSITNITGIIETANNKATDVGGGAYVKGTLTCEN 480
Qy 481 SHRLQFLKNSSDKQGGGIYGEDNITLSNLTKTLFQENTAKEEGGLFIKGTDKALMTG 540
Db 481 SHRLQFLKNSSDKQGGGIYGEDNITLSNLTKTLFQENTAKEEGGLFIKGTDKALMTG 540
Qy 541 LDSFCLINNTSEKHGGGAFVKEISQTTYSDETIPGTPVHGETVITGNKSTGGNGGV 600
Db 541 LDSFCLINNTSEKHGGGAFVKEISQTTYSDETIPGTPVHGETVITGNKSTGGNGGV 600
Qy 599 CTKRLALNSLQISLSGNSAENGGAHTCPDSPPTADTAEPAAASATSTPKSAP-VS 659
Db 599 CTKRLALNSLQISLSGNSAENGGAHTCPDSPPTADTAEPAAASATSTPKSAPVS 658
Qy 660 TALSTPSSSTVSSLTLLAASQAASPATSKETQDPNADTDLIDYVVDTTISKNTAKRG 719
Db 660 TALSTPSSSTVSSLTLLAASQAASPATSKETQDPNADTDLIDYVVDTTISKNTAKRG 718
Qy 720 GIYAKKAKMSRIDQUNISENATEIGGGICCKESLELDALVSLVTENLVKKEGGLHAK 779
Db 720 GIYAKKAKMSRIDQUNISENATEIGGGICCKESLELDALVSLVTENLVKKEGGLHAK 778
Qy 780 TVNINLKSFGSFSNNKANSSTGVATTASAPAAAASLQAAAAAPSPATPYSGVVG 839
Db 779 TVNINLKSFGSFSNNKANSSTGVATTASAPAAAASLQAAAAAPSPATPYSGVVG 838
Qy 840 GAIYGEKVTFCQSGCTCFQSGNQADNNPSSSLNVQGGAIYAKTSLSIGSDAGTSYIF 899
Db 839 GAIYGEKVTFCQSGCTCFQSGNQADNNPSSSLNVQGGAIYAKTSLSIGSDAGTSYIF 898
Qy 900 SGNVSSTGKSTQGTGAIAGAIYSPVTILNCATPATSNNNTASATPKTSSDSSGNSIKDT 959
Db 899 SGNVSSTGKSTQGTGAIAGAIYSPVTILNCATPATSNNNTASATPKTSSDSSGNSIKDT 958
Qy 960 ICGAIGTAGTALTLSGVSRESGNTADLGAIGTLANANTPSATSGSONSITEKITLNGSFI 1019
Db 959 ICGAIGTAGTALTLSGVSRESGNTADLGAIGTLANANTPSATSGSONSITEKITLNGSFI 1018
Qy 1020 FERNOANKRGAIYSPSVSIKGNITFNQNTSTHDGSAIYFTKDATIESLGSVLFETGNVNT 1079
Db 1019 FERNOANKRGAIYSPSVSIKGNITFNQNTSTHDGSAIYFTKDATIESLGSVLFETGNVNT 1078
Qy 1080 ATQASSATSGQNTNTANTGAALFGDPGTQSSQTDAILITLLASSGNITFSSNLSQNNQGD 1139
Db 1079 ATQASSATSGQNTNTANTGAALFGDPGTQSSQTDAILITLLASSGNITFSSNLSQNNQGD 1138
Qy 1140 TPASFCSTAGVVKLSLQAAKGTISFFDCVHTSTPKTKGTONVYETLIDINKENSNPYT 1199
Db 1139 TPASFCSTAGVVKLSLQAAKGTISFFDCVHTSTPKTKGTONVYETLIDINKENSNPYT 1198
Qy 1200 GTIVFSELHENKSYIPQNALHNGTLLVLEKTELHVVSFEQKEGSKLIMEPGAVLSNQ 1259
```

```
Db 1199 GTIVFSELHENKSYIPQNALHNGTLLVLEKTELHVVSFEQKEGSKLIMEPGAVLSNQ 1258
Qy 1260 IANGALAINGLTIDLSSMGTPOAGEIFSPPELRIVATSSASGGVSSSTPTPKRISA 1319
Db 1259 IANGALAINGLTIDLSSMGTPOAGEIFSPPELRIVATSSASGGVSSSTPTPKRISA 1318
Qy 1320 AVPSGSAATPTMSSENKVFELTGLDITLIDPNGFNQNPMLGSDLDVPLIKLPTNTSDVOVY 1379
Db 1319 AAPSGSAATPTMSSENKVFELTGLDITLIDPNGFNQNPMLGSDLDVPLIKLPTNTSDVOVY 1378
Qy 1380 DLTLSGDLFPQKGYMGWTWLTDSNPQTGKLQARWTFDTYRRWVYIPRDNHFYANSILGSQ 1439
Db 1379 DLTLSGDLFPQKGYMGWTWLTDSNPQTGKLQARWTFDTYRRWVYIPRDNHFYANSILGSQ 1438
Qy 1440 SMIVVYKQGLINMLNNAARFDDIAYNNFVWVGVIPLAQOQTPLEEFPSYISRGTSVAIDA 1499
Db 1439 SMIVVYKQGLINMLNNAARFDDIAYNNFVWVGVIPLAQOQTPLEEFPSYISRGTSVAIDA 1498
Qy 1500 KPRQDFILGAAFSIVGKTKAIKKMHNYFHKGSYVSQASVYGGKFLYFLNKHQHWALP 1559
Db 1499 KPRQDFILGAAFSIVGKTKAIKKMHNYFHKGSYVSQASVYGGKFLYFLNKHQHWALP 1558
Qy 1560 FLIQGVSVSYGHKHTTTLTPSYIHERNKGWEDLGWLADLRISMDLKEPSKDSKRITVY 1619
Db 1559 FLIQGVSVSYGHKHTTTLTPSYIHERNKGWEDLGWLADLRISMDLKEPSKDSKRITVY 1618
Qy 1620 GELEYSSIRQKQFTIEDYDPRHFDCCAYRNLSLPVGCAGEAIIANCNTILMYNKLALAYMP 1679
Db 1619 GELEYSSIRQKQFTIEDYDPRHFDCCAYRNLSLPVGCAGEAIIANCNTILMYNKLALAYMP 1678
Qy 1680 SIYRNPVCKYRVLSNNEAGOVICGVPTRTSARABYSTOLYLGPEWTLTYGNTYIDVGMYT 1739
Db 1679 SIYRNPVCKYRVLSNNEAGOVICGVPTRTSARABYSTOLYLGPEWTLTYGNTYIDVGMYT 1738
Qy 1740 LSQMTSCGARMIF 1752
Db 1739 LSQMTSCGARMIF 1751
RESULT 2
Q9PJY2 ID Q9PJY2 PRELIMINARY; PRT; 1672 AA.
AC Q9PJY2:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE POLYMORPHIC MEMBRANE PROTEIN B/C FAMILY.
GN TC0694.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Ewison M., Nelson W., DeBoy R., McClarty J., McClarty G., Salzberg S.L.,
RA Glavin J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002338; AAF39510.1; -.
DR TIGR; TC0694; -.
DR InterPro; IPR003357; -.
DR Pfam; PF02385; OMP; 1.
SQ SEQUENCE 1672 AA; 176295 MW; 4ABF190DA4DF8BD6 CRC64;
```

Query Match 72.6%; Score 6488.5; DB 2; Length 1672;  
Best Local Similarity 73.0%; Pred. No. 0;

Matches 1292; Conservative 145; Mismatches 214; Indels 119; Gaps 20;

```
QY 1 MKWLSATAFAAALPSVSGFCPEPEKELNFSRVETSSSTFTTETIGEAGAEYIVSGNASF 60
Db 4 MKWLSATAFAAALPSVSGFCPEPEKELNFSRGTGSSSTFTTETIGENGTEYIVSGNSFF 63
QY 61 TKFTNI-----PTTD---TTPPNSNSSSSGE-----TASVSDSDSTTTTPDPKGGGAFY 109
Db 64 TNPNTNPVKPPTDSDSTPTTSSAVDPTKEKIVRASSSPNSGSDTSATPDPKGGGAFY 123
QY 110 NAHSGVLSFWRTGSGTSLSEIKMTGEGGALFSGOGLLFTDLTSLITONNLSQLSGGA 169
Db 124 NEHSGILSPWARGVSGSLTSLNKMWTGGGALYSOGELLFTDLTGLTQCNLSQLSGGG 183
QY 170 IFGGSTISLGIKATFSCNSAEVPAPVKPTEPKAQTASSETSGSSSSGNDVSVSPSS 229
Db 184 IFGGSTISFGINOAFTSSNTAEV---VPEETPNPNPGTQTTSQ-----PSPTS- 231
QY 230 RAEPAAANLQSHFICATAPAAQTDTEETSPSHKPGSGGAIYAKGDLTTADSOEVLFSIN 289
Db 232 -----KVQSLFTYSSSTQANGSADGSDTPSHKPGSGGAIYATGDLTISDSQEIYFSVN 284
QY 290 KATKDGCAIFAEKDVSFENITSLKVOTNGAEKGGAIYAKGDLISIOSSKOSLFSNYSKQ 349
Db 285 KASKDGCAIFAEKWNFENITLKVQNGAEKGGIYASGDLISIOSSKOSLFSNYSKQ 344
QY 350 GGGALYVEGGINFQDLLEIRIKYNKAGTETKRTITLPSLKAQASAGNADAWASSPSQSGS 409
Db 345 GGGALYIEGNVDFKOLEIRIKYNKSGTETKVKVTLSPLEAQTINKSSVTA----- 394
QY 410 GATTVSDSGSDSGSDSIRSEVPVPTAKGGGLYTDKNLSITNITGIEIANKKATDVGGG 469
Db 395 -----ASQSGPNNTPTPTPVPTAKGGGLYTEKNLSITNITGIEITNNKATDVGGG 445
QY 470 AYVKGTLTENSRLQFLKNSDKGGGLYGEDNITLSNLTKTLFQENTAKEEGGLFI 529
Db 446 AYVKGTLTKDHRQLQFNSSKGGGLYEDTITLSNLTKTLFQENTAKEEGGLYI 505
QY 530 KGTDKALMTGLDLSFCLINNTSEKHGGGAFVTKETISQVTSQVETIPGTPVHGETVITG 589
Db 506 QGDDKTLTWTGLDLSFCLIDNTSATHGGGAVTKETISQVTSQVVEEPPGITPVHGETIISG 565
QY 590 NKSTGGNGGVCTKRLALNSLQISISIGNSAABNGGAHTCPDPSFTADTAEOPAASAA 649
Db 566 NKATGGSGGVCTKHLVLNLQITISSEFASENGGACTCPDNFP----- 611
QY 650 TSTPKAPYSTALSTPSSVSLTLLAASSQASPATSNKETODPNADTDLIDYVVDIT 709
Db 612 -----APTA---STPSTN-----QTAAPKDDKDFLIDYVVSFT 641
QY 710 ISKNTA-KRGGGIYAKKAKMSRIDQLNISSENSATEIGGGICCKESLELDALVSLSVTENL 768
Db 642 IDKNATKKGAGYAKKALSRIDELNISDNAQETGGGFCCTESLELDTIASLSVTKNL 701
QY 769 VKKEGGLHAKVTNINSLKGSFSSNNKANSSTGVATTASAPAAAASLQAAAAAPSS 828
Db 702 AKKEGGLHAKTLINISLKSGLSFSNNTANSSTGVATTATTSQSPTVSFSLPRATAGSS 761
QY 829 PA---TPTYGVVGGAIGEKVTFSCQSGTCQFSCNOAIDNNPSSSLNVQGGAIYAKT 884
Db 762 PAPAQTPTYAGVVGGAIGETVFSFKSGLCQFTENSADINTPSSPSLNVQGGAIYAKT 821
QY 885 SLISGSDAGTSYIFSGNSVSTGKSQTTGQIAGGAIYSPTVTLNCPATFSNNTASTATPK 944
Db 822 SLISAEDEPSTSVFSKNSVSGKATQTTQIAGGAIYSPTVTLNCPQTVFSGNSASWAT-- 879
QY 945 TSSEDGSSGNSIKDTIGGAAGTATLGSVRFSGNLTADLGAAGTGLANANTPSATSGSO 1004
Db 880 ---TNPPTSCTSPKDTIGGAAGTTISLSKTSHFSENTADLGAAGTGL-----SGGSS 928
QY 1005 NSITEKITLNGSFTFERQNAKRGAIYSPSVISIKGNNTFNQNTSHDGSALYFTKDAT 1064
Db 929 SNLTKITLNGSFTFERKNAKRGVIYAPSVISIKGNNTFNQNTSHDGSALYFTKDAT 988
```

```
QY 1065 IESLGSVLFTGNVNTATQASS-ATSGQNTNTANYGAAIFGDPGTTQSSOTDAILTLASS 1123
Db 989 IESLGSVLFTGNVNTAEQASSTATGQTTNTNTYGAAGFDPGCT---QTDDTTLKLAIASS 1045
QY 1124 GNTTFNSNLONQOGDTPASKFCISAGYVKLSLOAKGKTISFFDCVHTSTKKTGTSQNV 1183
Db 1046 GNTTFNSNS-ONTATNPATKFCISISGYVKLTLOAAQGTISFFDSIRTSTKKTGQNS 1104
QY 1184 YETLDLNKENSNPYTGTVFSSSELHENSYPQNALHNGTLLVLEKTEHLVSVFQKE 1243
Db 1105 YETLDLNKENSNTYAGTVLFSSSELHEVKSYPQNVVHNGTLLVLEKTEHLVSVFQKE 1164
QY 1244 GSKLIMEPCAVLSNQNIANGALAINGLTIDLSSMGTPQAGEIEFSPPELRIVATSSAGG 1303
Db 1165 GSKLIMEPCAVLSNQNIANGALAINGLTIDLSSLGAPQTGEVSPPELRIVATSSAGG 1224
QY 1304 SGVSSSIPNPKRISAAVPSGSAATPTTSENKVFITGDLTLIDPNGNFQNPMLGSDL- 1362
Db 1225 GVGYGTV-TASKNLSAASPT-VAATNPTWADNKVFTGALTLDIDPNGNFYQNPILGTDLT 1282
QY 1363 DVPILKLPNTSDVOYDVLTLSGDLFPQKGYMGTWTLDSNPQTGKLQARWTFDYRIHWY 1422
Db 1283 DVPILKLPNTANGVDVSNLTLSGDLSPKGYGTGTWTLNPDQTKGVVANKKFDYRIHWY 1342
QY 1423 IPRDNHFANSILGSONSMIVVKQGLINNMNNAFDDIAYNNFWVSGVGTFLAQOCTPL 1482
Db 1343 IPRDNHFANSILGSONSMIVVKQGLINNMNNAFDDIAYNNFWVSGVGTFLAQOCTPL 1402
QY 1483 SEEFYSYSGTSAIDAKPRQDFILGAAPSKIVKTKAKIKMHNYFHKSEYSAQAEVYG 1542
Db 1403 SEEFYSYSGTSAIDAKPRPDPILGAAPSKVMVGRTKAIVKHNYSHKSEYSAQAEVYG 1462
QY 1543 GKFLYELLANKHGWALPELLOQVGVSYGHKHDFTTLYPSIHERNKGDWEDLGLADLRIS 1602
Db 1463 GKFLYELLANKHGWALPELLOQVGVSYGHKHDFTTLYPSIHERNKGDWEDLGLADLRIS 1522
QY 1603 MDLKEPSKDSKRRITVYGELEYSSIRQKQFTEIDYDPRHFDCCAYRNLSLPVGCAGEGAI 1662
Db 1523 MDVKEPSKDSKRRVYGELEYSSIRQKSFTEIDYDPRHFDCCAYRNLSIPMGCYFEGAI 1582
QY 1663 MNCNIMLNKALAYNPSTYRNPNCKYRVLSNAGOVICGVPTTTSARAESYQIYLG 1722
Db 1583 MSYDIILMKNLSLAYNPSTYRNPNCKYRVLSNAGOVICGVPTTTSARAESYQIYLG 1642
QY 1723 PFWTLVGNVTIDVGMTLSQMTSCGARMIF 1752
Db 1643 PFWTLVGNVTIDVGMTLSQMTSCGARMIF 1672
```

```
RESULT 3
084419 PRELIMINARY; PRT; 1770 AA.
AC 084419;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE PUTATIVE OUTER MEMBRANE PROTEIN C.
GN PMPC.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/OW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
DR EMBL; AB001315; AAC68011.1; -.
```



[illegible]







Db 1205 -TGCTILSELHENKSIYIPQKVTFAHGNLILGKNABLSVVSFTQSPGTITMGPGSVLSN 1263  
QY 1258 QNIANGALAINGITLIDLSMGTPQAGEIESPPELRIVATISSASGSGVSSSIPTNPKRI 1317  
Db 1264 HSKAAGGAIINNVIIIDFSEIVPTKONATVAPTPLKLVSTNAD----- 1306  
QY 1318 SAAVPSGSAATPTMTSENKVFITGDLTLDNPGNFFONPMLGSDLDVPLIKLPTNTSD-V 1376  
Db 1307 -----SKDKIDITGTVTLLDPNGLYQNSYLGEDRDITLFDNIDNSAGAV 1351  
QY 1377 QVYDLTSLGDLPPQGYMCTWTWLDSPQTKLQARWTEDTYRWVYIPRONHYANSILG 1436  
Db 1352 TATNVTLQGNLGAKKGLTGNLDPNNSGSKILKWTFDKYLRPWYIPRDNHYINSIG 1411  
QY 1437 SONSMIVKQGLINMNNARFDDIAYNFWVSGVGTFLAQOQTPLSSEFYSYRGTSTA 1496  
Db 1412 AQNSLVTVKQGLGNMNNAREDFAFNFWASIGSFLRKEVSRNSDSFTVHGROYTAA 1471  
QY 1497 IDAKPRQDILGAASKIVGKYATKMKMNYFHKHSEYQASVYGGKFLYLLNKHQGW 1556  
Db 1472 VDAKPRQEPILGAASFQVFGHAESEYHLDNYKHKGSGHSTQASLYAGNIFYFPAIRSR-- 1529  
QY 1557 ALPELQGVVYGHKTHDHTTLYPSYHERNKGDWEDGLWDLRLISMOLKEPSKSKRI 1616  
Db 1530 --PILFQGVATYGYQWHDFTTYTPSYEEKNMANNWDSIAWFLDLRFSVDLKEPOPHSTARL 1587  
QY 1617 TVGYELEYSSIRQKOFTEIDYDPRHDDCAYRNLSLPVGCAGEAIGMNCNIMLYNKLALA 1676  
Db 1588 TTYTEAYTRIQRKTELDYDPRFSACSGLNLAIPTGFSVDGALAWREIILYNKVSAA 1647  
QY 1677 YNPSTYRNPNVCKYRVLSNEAGQVTCGVPTRTSARAERYSTOLYGPFWTLXGNYTIDVG 1736  
Db 1648 YLPVLIRNPNKATYEVLSYKKEGNNVWVLPTRNAAREVSSQIYLGSYWTLGYTYTIDAS 1707  
QY 1737 MYTLOMTSCGARMIF 1752  
Db 1708 MNTLVQMANGIRFVF 1723

RESULT 6  
Q9RB59  
ID Q9RB59 PRELIMINARY; PRT: 1723 AA.  
AC Q9RB59;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE POLYMORPHIC MEMBRANE PROTEIN B FAMILY.  
GN PMP\_20  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J138;  
RA Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,  
RA Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,  
RA Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,  
RA Ishii K., Shiba T., Hattori M., Kuhara S.,  
RT "Comparison of outer membrane protein genes omp and pmp in the whole  
RT genome sequences of Chlamydia pneumoniae isolates from Japan and US."  
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10871362;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.,  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
RT from Japan and CWL029 from USA."  
RL Nucleic Acids Res. 28:2311-2314(2000).  
DR EMBL; AB033817; BAA85968.1; -.  
DR EMBL; AP002547; BAA98746.1; -.  
DR InterPro; IPR001254; -.

DR InterPro; IPR003357; -.  
DR InterPro; IPR003368; -.  
DR Pfam; PF02385; OMP; 1.  
DR Pfam; PF02415; DUF145; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
SQ SEQUENCE 1723 AA; 179593 MW; 4507D5B8A0E3EF01 CRC64;

Query Match 34.8%; Score 3111.5; DB 2; Length 1723;  
Best Local Similarity 37.6%; Pred. No. 1.8e-157;  
Matches 727; Conservative 284; Mismatches 528; Indels 397; Gaps 45;

QY 1 MKWLSATAVFAAALPVSVSGFCFPEPEKELNFSRVETSSSTT-----FTETIGEA-GAEYI 53  
Db 1 MKWLPAFAVFAAALPVALTAFGDPAASVEISTSHTGSGDPTSDAALTGTOSSTEDGTTVT 60  
QY 54 VSGNASFTFTNPTTDTTPTNSNS-----SSSGET----- 86  
Db 61 IVGDITFTFTNIPVPVPTDANDSSSSKGGSSSGATSLIRSNLHSDFDFTKSVL 120  
QY 87 -----ASVSDSDSTTTTTPDPKGGGAFY-NAHSG 114  
Db 121 DLYHLFPASNTLNPALLSSSSGSSSSSSSSGSAVVAADPKGGAAYSEANG 180  
QY 115 VLSFMTSRGTEGLTSEIKMTGEGGAIFSQGELLEFLDTLSTLQNLNLSQSGAIF--- 171  
Db 181 TLTFTTDSNGPSLTLQNLKMTGDGAAYSKGPLVFTGLKLNLTFTGNESQSGGAAYTEG 240  
QY 172 -----GG-----STIS 177  
Db 241 ALTQAIQEAIVTGTNTSAGQGGAIYVKEATLFLNALDSLKFEKNTSGQAGGIIYESTLT 300  
QY 178 LSGITKA-TFSCNSARVPAPVKPTPEKQAQTASETSGSSSSGSDSVSPSSSRAPFAA 236  
Db 301 ISNITKSIIEFISNKASVPAPAPEPTSP-----APSSLINSTTIDTTLQTRAASATPAVA 355  
QY 237 NLQSHFICATAPAAQTDTTETSTPSHKPGSGGAIYAKGDLTI----- 278  
Db 356 PV-----AAVTPPTISTQETA-----GNGGAIYAKQGISISTFKDLTFKSNASVDAT 403  
QY 279 -----ADSQEV-----LFSINKATDGGGAIFAEDKVDVSENIISLVK 314  
Db 404 LTVDSSTIGSGGAIFAADSIQIQCTGTTLFSGNTANKSGGGIYAVGQVTLIEDIANLKM 463  
QY 315 QNGAEEKGAIYAKGDLSTQSSK-QSLFNSNYSKGGALYVEGGINFODLEIRIKYN 373  
Db 464 TNNCKEGEGAIYTKKALINNGAILITFTSGNTSTONGGAIFAVGGITLSDLVEVEFSKN 523  
QY 374 KAGTFETKKTITLPSLKAQASAGNADAWASSSQSGSGATTVSDSGSSGSDSDTSETVP 433  
Db 524 KTGNY-----SAPITKA---ASNTAPVWSSTTAAAPV-----PAAAAAP 561  
QY 434 VT--AKGGGLYTDKNLSITNITGIIIEIANNKATDVGCGGAYVKGTLTCENSHRLOFLKNS 491  
Db 562 VTNAAKGGLYSTEGLTVSGITSLFENNECONQGGAYVTKTFCQSDSHRLOFTSNKA 621  
QY 492 DKQGGIYGEDNTLNLTKTLFOENTKAKEGGGFIKGTDKALMTGLDSCFLNNTS 551  
Db 622 ADEGGGLYCGDDVTLNLTKTLFOENSEKKGGLSL-ASGSLTWTLSLESCFLNANTA 680  
QY 552 EKHGGGAFVTKESQTYT---SDVETIPGTPVHGETVITGNKSTGGNGGGVCTKRLALS 608  
Db 681 KENGGGANVPENIVLTFTTPTNEPAPVQPPVYGEALVTGNTAT-KSGGGIYTKNAAFS 739  
QY 609 NQOSISISNSAENGGAHTCPDPTADTADQPAASAASTSTPKSAPVSTALSTPSS 668  
Db 740 NLSSVTFDQNTSSENG-----ALLTQKAAD----- 765  
QY 669 TVSSLLTLLAASSQASPATSNKETQDPNADTDLIDYVVDTTISKNTAK-KGGGIYAKKAK 727  
Db 766 -----KTDCSFTYITNVNNTNTATNGNGGIAGGAKAH 797  
QY 728 MSRIDQLNISENSATEIGGICCKEISLELDALVSLSVTENLVKGGGLHAKTVNISHLK 787

Db 798 FDRIDNLTVQSNQAKK-GGGVLEDALILEKVITGVSQNTATESGGGIYAKDIQLQALP 856  
Qy 798 SGFSFNKANKSSSTGVATTASAPAAAAASLQAAAAAPSPAPPTYSGVVGGAIYGE-K 846  
Db 857 GSFTITDNKTSLSL-----TSNLYGGGIIYSSGA 886  
Qy 847 VTFSCSGTCFSGNQADNNPSSSLNVQGGAIYAKTSLSIGSSDAGTSIFSGNSYST 906  
Db 887 VLTNISGTFGITGNSVINTATSQDA-DIQGGIYATTSLSI--NQCWTFILFSNNSAAT 943  
Qy 907 GKSQTTQGIAGGIYSPVTVL---NCPATFSNNTASIAATPKTSSDSSGNSIKDTIGGA 963  
Db 944 KKTSTTKIAGGAISAAVTTIENNSOPIIFLNSAK----SEATTAATAGN-KDSCGGA 997  
Qy 964 IAGTAITLSGYSR--FSNTADLGAAG--TLANANTPSATSGNSQNTSEIKITLNGSFI 1019  
Db 998 IAAANSVTLTNNPETFKGNYAETGAICIDLITNGSPPRKVS-----IADNGSVL 1047  
Qy 1020 FERNOA-NKRGAIYSPSVISIKGNITFNQNTSTHDGSAIYFTKDATIESLGSVLFTGNV 1078  
Db 1048 FQDNSALNRGAIGETIDISRTGATFGNSKDKGSAICCTALTALPNSQLIFENKV 1107  
Qy 1079 TATQASSATSGONTWANYGAIFGDPGOTQOSSODAILTLASSGNTTFFSNLSQNNQ 1138  
Db 1108 TETTATTAKS-----INNLAGAIYGNNET-----SDITISLSAENGSTFFKNLNC----- 1152  
Qy 1139 DTPASKFCISAGYVKL-SLOAKGKTIISFFDCVHTSTKTKTSTQNTVYTLINKEENSNP 1197  
Db 1153 -TATNKYCSIAGNKFTFAIEASAKAISFYDAVNVSTKETNA-----OELKNEKATS-- 1204  
Qy 1198 YGTIVFSSELENKSYIPQNAIHLNGLVLEKTELHVWFEQKESKLIMEPGLVLSN 1257  
Db 1205 -TGILFSGELHENKSYIPQVTAHGNLILGKNAELSVSFTQSPGTTIMPGSVLSN 1263  
Qy 1258 QNIANGALAINGLTIDLSMGTPOAGETSPPELRIVATTSASGGSGVSSSIPNPKRI 1317  
Db 1264 HSEKAGGIAINNVITDFSEIYPTKDNATVAPPTLKLVSRTNAD----- 1306  
Qy 1318 SAAVPSGSAATPTMSENKVFELTDLIDPNGFYQNPMLGSLDVLPLKLPNTWSD-V 1376  
Db 1307 -----SKDKIDITGVTLIDPNGNLYQNSYLGEDRDITLFDNIDSASGAV 1351  
Qy 1377 QVYDITLSDGLFPQKGYMGWTLDNSNPOTGKLOARWTFDYRRWVYIPRONHFYANSTLG 1436  
Db 1352 TATNVTILOGNLGAKKYLGTLNLDPNSSGSKIILKWTDFKYLWPYIPRONHFIYNSIW 1411  
Qy 1437 SONSMIVVQGLINMLNARFDDIAYNFWVSGVGTPLAQOQTPLSEFSYYSRGTSVA 1496  
Db 1412 AQNSLVTYKQILGNMLNARFEDPAFNFWASAIKGSPLRKEVSRNSDSFTYHGRGYTAA 1471  
Qy 1497 IDAKPRODFILGAAFSKIYKTKAIKMHNYFHKGSSEYYSQASVYGGKFLYELLNKHGW 1556  
Db 1472 VDAKPRQFELGAASFQVGHAESEYHLDNTYKHGSGHSTQASLYAGNIYFFPAIRSR-- 1529  
Qy 1557 ALPELIQVWYSGIHKDHTTYLPSIHERNGDMEDLGLWADLRISMOLKEPKSKDSKKRI 1616  
Db 1530 --PILFQGVATYGYMQHDTTYTYSIEKNNANWDSIAWLFDRFSVDLKEPQPHSTARL 1587  
Qy 1617 TVYGELEYSSIRQOFTIEDYDPRHFDCAVRNLSLPGVCAVEGAIMCNILMYNKLALA 1676  
Db 1588 TPTYEAETRIROEKFTEDLDYDPSFSACYSYCNLAIPTFGSDGALAWREILYKNVSA 1647  
Qy 1677 YMPSTYRNPNVCKYVLSSNAGOVICGVPRTTSARAETSQLYLGPWPWLYGNVYTDVG 1736  
Db 1648 YLPVILRNPNKATYEVLTSEKKGNNVNVLPTRNARAEEVSSGIYLGSTWLYGTITDAS 1707  
Qy 1737 MYTQSMTSCGARMIF 1752  
Db 1708 MNTLVQAMANGIRFEV 1723

RESULT 7

Q9K2C1  
ID O9K2C1 PRELIMINARY; PRT; 1732 AA.  
AC O9K2C1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE POLYMORPHIC MEMBRANE PROTEIN B/C FAMILY.  
GN CP0212.  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hacke E.K., Peterson J., Umayam L.A., Utterback T.,  
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,  
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,  
RA McClarty G., Salzberg S.L., Eissen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
pneumoniae AR39.";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
DR EMBL; AE002181; AAF38082.1; -;  
DR TIGR; CP0212; -;  
DR InterPro; IPR001254; -;  
DR InterPro; IPR003357; -;  
DR Pfam; PF02385; OMP; 1;  
DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
SQ SEQUENCE 1732 AA; 180631 MW; DB02D918C777B525 CRC64;

Query Match 34.8% Score 3111.5 DB 2; Length 1732;

Best Local Similarity 37.68; Pred. No. 1.8e-157; Indels 397; Gaps 45;  
Matches 727; Conservative 284; Mismatches 528;

Qy 1 MKWLSATAVFAAALVPSVSGFCFPPKELNFRVETSSSTT-----FTETICEA-GAEYI 53  
Db 10 MKWLPATAVFAAALVPSVSGFCFPPKELNFRVETSSSTT-----FTETICEA-GAEYI 53  
Qy 54 VSGNASFTKFTNIPTDTTPTNSNS-----SSSGSET----- 86  
Db 70 IVGDITFTFTNIPVPVTPDANDSSSSSKGGSSSGATSLIRSSNLHSDFDFTKDSVL 129  
Qy 87 -----ASVSESDSTTTTDPKGGGAFY-NAHS 114  
Db 130 DLYHLFFPSASNTLNPALLSSSSSGSSSSSSSSSSSSSASAVVAADPKGAFAFTNEANG 189  
Qy 115 VLSFTMRSGTEGLSLSEIKMTGEGGAIFSGQCELLFTDLTSLITQNNLSQLSGGAIF--- 171  
Db 190 TLTFTTDSGNPGLTLQNLKWTGDGAAIYSGPLVFTGLKMLTFTGNSQKSGGAAYTEG 249  
Qy 172 -----GG-----STIS 177  
Db 250 ALTTQAIWEAVTFTCNTSAGCGGAIYVKEATLFLNALDSLKEFKNTSGOAGGIYTESTLT 309  
Qy 178 LSGITKA-TFCNSAEVPAPYKPKPEKAQTASFTSGSSSSGNDSSVSSSSRAEPAAA 236  
Db 310 ISNTKISIEFTSNKASVPAPEPTSP-----AFSSLINSTIDTSTLTQRAASATPAVA 364  
Qy 237 NLSHFTCATAPAAQTDSTSTPSHKPGSGGAIYAKGDLTAI----- 278  
Db 365 PV-----AAVTPPISTQETA-----GNGGAIYAKGISISTFKDLTFKNSASVDAT 412  
Qy 279 -----ADQEV-----LFSINKATKOGGAIFAEDKVSFENITSUKV 314  
Db 413 LTVDSSTIGESGAIFAADSIQOCTGTLTSGNTANKSGGIYAVGQVLTEDIANLKM 472  
Qy 315 QTNGAEKGGAIYAKGDLDSIOSK-QSLFNSNYKQGGGAIYVEGGINFDLEEIRIYN 373  
Db 473 TNNTCKGEGGAIYTKKALTINNGALLITFSGNTSTDNCGAIFAVGGITLSDLVVRFSKN 532  
Qy 374 KAGTFTETKKITLPSLKAQASAGNADAWASSPQSGSGATTVSDSGSSGSDSTETVP 433



Db 172 AVSAYSIDIAKTTTAAALDQNTSTK-----NGGALCSTANTTVOGNSGTVTFSSNTATDK 226  
QY 295 GGAIFA-EKDVFNITSLKVOTNGAEKGGAIYAKGDLISQSSKSLFNSNYSKQGGGA 353  
Db 227 GGGIYSKEKDSLTANTGVVTFKSTAKTGGAWSDDNALALTQNTQVLFQEN----- 278  
QY 354 LYVEGGINFQDLEERIKYIKNAGTPTETKITLPSLKAQASAGNADAWASSPQSGGA-- 411  
Db 279 -----KTTGSAQA-----NNPEGCGGAIC 298  
QY 412 ---TTVSDS---GSSSSGSDSTSETVPVAKGGGLYTDKNLSITNITGIEIANKATD 465  
Db 299 CYLATATOKTGLAISQNMESFTST--TTANGGAIYATK----- 336  
QY 466 VGGGAYVKGTLTCENSHRLQFLKNSSDKQ-GGGIYGE-DNITLSNLTGKTLFQENTAKEE 523  
Db 337 -----CTLDGNTTTLTFOQNTATAGCGGAIYETEDFSLKSGTGVTFSTNTAK-T 385  
QY 524 GGGLFIKGTDKA-----LTMTGLDSFCLINNTSEKHGGGAFVTKETISQTYTSDVETIPQI 578  
Db 386 GGALYSKGNLSLTGNTNLLFSGNKATGPNSSANQEGCGGAILAFIDSGVS DK--TGL 442  
QY 579 TPVHGSETV-ITONKSTGGNGGVCVKRLALSNLQISISGNSAENGGAHTCPDSFPPTA 637  
Db 443 STANNQEVSLTSNAAT-VSGGAIYATKCTLTGNGSLTFDGTAGTSGGAIYETEDFTL- 500  
QY 638 DTAEQPAAASAATSTPKSAPVSTALSTPSSSTVSSITLLAASSQASPATSNKETODPNAD 697  
Db 501 -----TGSTGV----- 507  
QY 698 TDLIDYVVDVTTISKNTAKGGIYAKKAKMSRIDQLNISENATEIGGICCKESLEID 757  
Db 508 -----TFSTNTAKTGALYSK-----GNN----- 526  
QY 758 ALVSLSVTENLVKGGGLHAKTVNISNLKSGFSSNNKANSSTGTGAVTASAPAAAAAS 817  
Db 527 ---SLSGNTNL-----FSGNKATGPNSSANQEGCGGAILISF 561  
QY 818 LQAAAAAPPSPATPTYSGVVGAIYGEKVTFSCGTCQPSGNOAIDNNPQSSINVOG 877  
Db 562 LESASVST-----KKGLWIENENVLSGNTA-----TVSG 592  
QY 878 GAIYAKTSLISIGSDAGTSYIFSGNSVSTGKSQTTGQIAGGAIYSPPTVTLNC-----PAT 932  
Db 593 GAIYATKCALHCN---TTLTDEGN-----TAETAGGAIYETEDFTLTGSGTGT 639  
QY 933 FSNNTASTATPKTSEDGSGSINSIDTGGATAGTAITLISGVSRFSGNTADLGAIGHIA 992  
Db 640 FSTNTA-----KTAGALHTKGN-----TSFTKNKALVFSGNSA-----TA 674  
QY 993 NANTPATSATSGQNSI-----TEKITL-ENGSTIFERNOANKR-GAIYSPSVSIKG 1040  
Db 675 TATTTDQEGCGGAILCNLISESDIATKSLTLENESLSPINNTARSGGIIYAPKCVISG 734  
QY 1041 N-NITFNQNTSHDGSIAIYFTKDATIESLGSVLYFTGNNVNTATQASSATSGQNTANTYGA 1099  
Db 735 SESINFNGNTAETSGGAIY-SKNLSITANGPVSTNN-----SGGKG-----GA 777  
QY 1100 AIFGDPGTTQSSQTDAILTLTLASSGNITFSSNNLQNNQGDTPASKFCSTAGVVKLSLQAA 1159  
Db 778 IYIADSGE-----LSLEAIDGDTFSGN--RATEGTSPNSIHLAGAKAITKLAAA 826  
QY 1160 KGKTSFFDCVHTSTKTKSTQNVYETIDIN-----KEEN-----SN 1196  
Db 827 PGHTIYFYDPITMEAPASGTT---TEELVINPVKAIYPPPOPKNGPTASVPVVPVAPAN 883  
QY 1197 PYTGTVFSS-ELHENKSIYQNAI-----LHNGTLVLKTEKTLHVVSFEQKESGKL 1247  
Db 884 PNTGTVFSSGKLPQSDASIPANTTTILNQKINLAGGNVVLKEGATLQVYSFTQPDSTV 943  
QY 1248 INEPGAVL--SNQNTANGALNGLTIDLSMGTPQAGEIFSPPELRIVATSSASGSG 1305  
Db 944 FMDAGTTLTETTTNTDGSIDLNLSNLDALDGKR-----MITIAVNSTSGGLK 993

QY 1306 VSSSIPTNPKRISAAPVSGSAATTPTMSENKVFLTGDLTLIDPNGNFYQNPMGLSDLDVP 1365  
Db 994 IS-----GDLKFHNNEGSFYDNPGLKANLNP 1020  
QY 1366 LIKLTNTSDVOVYDLT--LSGDLFPQKGYMGTTWLDNSNPQTG-----KLQARWTFDTYR 1418  
Db 1021 FUDLSSTSGTGNLDDFNPIPSMAAPDYQGSWTL--VPKVGAGGKVTVAEWOALGYT 1078  
QY 1419 -----RWYIIPRDNHFYANSILGSONSMIVYKQGLNNLNNAAREDDIAYNFWVSGVGT 1473  
Db 1079 KPPELRATLVP-----NSLWNAYVNIHSIQOETA-----TAMSDAPSHPGITGGIGN 1126  
QY 1474 FLAQCCTPLSEEFSSYSGTSAIDAKPRODFILGAATSKIVGKTK--AIKMHNVFHKG 1531  
Db 1127 AFHQDKQENAGFRILSRGIYVGGSMTPPQETFAVAFSOLFGKSKDYVVDIKSQVYAG 1186  
QY 1532 S---EYSQASVYGG---KFLYLLNKQHGWAFLPIOGVSVSYGHKKHDTTLYPSIHER 1585  
Db 1187 SLCAQSSYVIVPLHSSLRHRVLSKVLPELPG-ETPLVLHGQVSYGRNHHNMTTKLAN-NTQ 1244  
QY 1586 NKGDEDLGWLADLRISMDLKEPSDKSKRITVYG---ELEYSSIROKQFTIDYDPRHF 1642  
Db 1245 GKSDWDHSFAVEVGSL---PVDLNYRLTSYSPYVKLQVSVNKGQFQEVAAADPRIF 1300  
QY 1643 DDCAYRNLSLPGVCAV--EGALMNCNIMYNKALAYMPSIYRNNPVCYRVLVSNEAQ 1700  
Db 1301 DASHLVNVSIPMLGTLFKHESAKPPSALL---LTGAYDAVDRDHPHC-LTSLTNGTWS 1355  
QY 1701 VICGVPTRTSARAESTOLYLGPFWTLYGNYTIDVGMVTLTSLQMTSCGARMIF 1752  
Db 1356 TFAATNLSROAFPAEASGHLKLLHGLDCAFGSCELSSSSSYNANCGRYSF 1407

## RESULT 9

Q9JRW2 ID Q9JRW2 PRELIMINARY: PRT: 1276 AA.  
AC Q9JRW2:  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE POLYMORPHIC OUTER MEMBRANE PROTEIN G/I FAMILY (POLYMORPHIC MEMBRANE  
GN PMP\_6 OR CP0309.  
OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=835556;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10871362;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
from Japan and CWL029 from USA."  
RL Nucleic Acids Res. 28:2311-2314 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
Berry O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,  
Werry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,  
Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,  
McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
pneumoniae AR39."  
RL Nucleic Acids Res. 28:1397-1406 (2000).  
DR EMBL: AP002546; BAA98652.1; --  
DR EMBL: AE002193; AAF38166.1; --  
DR TIGR: CP0309; --  
DR InterPro: IPR003357; --  
DR InterPro: IPR003368; --



SQL	SEQUENCE	1609 AA;	170883 MW;	7666C20F09C517B2	CRC64;
Query Match					
Best Local Similarity		8.3%; Score 743.5;		DB 2; Length 1609;	
Matches 406; Conservative 265; Mismatches 687;		Indels 459; Gaps 79;			
Qy	165	LSGGAIFGSGTISLS----	GITKATFCSNSA----	EVPAVPKKPTEP----	KAOTASE-----
Db	23	LSAGTAFSAHLSHSELDLVGNKQFESHA	VEAQTSLVKGSDPVNPSQKSEK	VLYT	82
Qy	211	-----TSGSSSS-----	GND-----	SVSPSSSSRAEP	233
Db	83	QVPLTQGGSGESLDADANFLEHFOHLP	FEETVFGIDOKLVMSDLDRNF	SOPTQ----	Ep 139
Qy	234	AAANLOSIFICATAPQAQTTEITSPSHK	PG-----	SGAIYAKCDLTIADSOEVLFSIN	289
Db	140	DTSNAVSEKI--SSDTKENRKDLETD	DPKSKGLKEVSDLPKSPETA	VAAAISEDLTSEN	198
Qy	230	KATKD--GGAIF-----	AEKDVSEFNI-----	TSKLVOTNGAEKGGAIYAK	329
Db	199	ISARDPLOGLAFFYKNTSSQISSEKDS	SFOGIIIFSGGANSGLGFENL	KAPKSGAAVYSD	258
Qy	330	GDLSTQSSQSL--PNSNYSKOGGALY	VEGGINF-----	ODLEEIRI--	370
Db	259	RDIIVFENLVKGLSFISCESLEDGSA	-----	AGVNIIVVTHCGDVTLTDCATGLD	LEALRLVK 314
Qy	371	KYNKAGTETKK-----	ITLPSLKAQASAGNADAWASS	SPQSGGATTYSDSCDSSG	423
Db	315	DFSRGAVFTARNHEVQNNLGAGLIS	VVGNKGAIVVEKNSAEKSGAF	ACGFSFYISNNE	374
Qy	424	SDSDTSEVPVTA	KGGLYTOKNLISNITITGIEI	ANKA-----	TD-VGGGAY- 471
Db	375	NTALWKENAQLS--GGAISSASD	IDIQNCSAIEFSGNQSLIAL	GEHMLDFVGGGALA	432
Qy	472	VKGLTCTENSHRLOFLKNSQKGGI	YGEDNITLSNITKGLTFOENT	AKBEEGGLFKG	531
Db	433	AQGLTLRNNAVOCVKNTSKTHGAILA	-GTVDLNETISEVAFKQNTAALT	TGGAL--SA	489
Qy	532	TDKALTMGLDSFCLINNTSEKHG	---GAFVYKEISQVTS	SDVETIPGTPVHGEIVI	587
Db	490	NDKVLIANNFGEIIFQNEVRHNG	GAIVYCGCRSNPKLEQKDS	GENINIGNS--	GAITF 546
Qy	588	TGNKST-----	GNGGCVCTKRLAL--SNLQ	SISISNSAAEN-----	GGGAHT 629
Db	547	LKNRASVLEVMQAE	DYAGGALGCHNVLLDSNS	GNIOFIGNIGSTWIGEVGGAIL	606
Qy	630	CPDSEPTADTA-----	EOPAAASATSTPKSAPV	STALSTPSSSTVSSLTLLA	ASSQ 681
Db	607	STDRTVTSNNSGDVVFKNGKQC	LQAKYVAPQETAPVESDASS--	TNKDEKSLINACSHGDH	665
Qy	682	ASPATSKRETODPNADTDL	LIDYVVDITISKNTAKKGGI	YAKKAKWSRIDOLNISENSA	741
Db	666	YPPKTVEE-----	VPPSLLEHPV-----	VSSTDINGGAILAQ-----	HIFITDNG 709
Qy	742	TEIGGGICCKESLELDALV	SLVTENLVKGGGGIHKATV--	NISNLKSGFGFSNNKAN--	S 799
Db	710	-----NLRFSGNL-----	GGGEESTVGD	LAIVGGALLSTNEVNV	C 746
Qy	800	SSTGVATTASAPAAAAA	SLQAAAAAPSSPATPTY	SGVVGGAIIYGEKVTF	SQCSGTQOFS 859
Db	747	SNQNVFSDNVTSGCDS-----	-----	GGAILAKKV-----	DIS 776
Qy	860	GNOAID--NPPSQSLNVQ	GGAIVAKTSLSIGSSD	AGTSYIFSGNSVSTGKS	QOTGOIAG 917
Db	777	ANHSVEFVNSGKFP-----	GGAVCALNE--SVNITD-----	NGSAVSFSKNRT--	RLGG 822
Qy	918	GAISPTVTLNCPATFSSNNTAS	IATPKT---SSEDSGSSG	SIKDTIGGAIAGTATL---	971
Db	823	AGVAAPOGSSV---TICNQ	GNIAFKENFVFGSEN	QNRSGG-----	GAIANSVNTODN 872
Qy	972	-SGVSRFSGNTADLGAAT--	GTLANANTPSATSG	SONSITEKITLENGS	FTIFERNQANKR 1028

Db	873	AGDILFVSNSTGSGGAI	FVGLSV-----	ASEGS--NPRTLIT	CNSGDILFAKNSTQTA 925			
QY	1029	GAIIYSPSVSIKGNIT	FNQNTSTHDGSAIYFT	KDATIESLGSVLFT	GNVNTATQASSATS 1088			
Db	926	ASL-----	SEKDSFGGAIYTONL	KIVKAGNVSFYGNR-----	APS 962			
QY	1089	QONTNANTYGA	AIFCDPCTTQSSQTDAILTL	LASSGNITFNNSLNQ	NOGQDTPASKFCSI 1148			
Db	963	CAGVQIA-----	DGGT-----	VCLEAFGDDIL	FEGNI--NPDGGSFNAIHLGCN 1003			
QY	1149	AGYVKLSLQA	AKGTISFFDCVHTSTKTKGT	QNVYE-----	TLIDN---KEENSN 1196			
Db	1004	DSKI--VELSAVDK	NIIFQDAI---TYE	ENTIRGLPKD	VSPLSAPSLIFNSKQDDSAQ 1059			
QY	1197	PYTGTVFSSSEL	HENKSYIPQNAIILHNG	ITIVLKEKTELHVVS	FEEQESKILMBPGAVL- 1255			
Db	1060	HHEGTIRFS----	RGVSKIPQIAA	IQEGTLALSONAELWAGL	KQETGSSIVLSAGSILR 1115			
QY	1256	-----	SNQNIANGALAIN-----	-----	GLTIDLSSMG 1278			
Db	1116	IFDSQVDSSAPL	PTEKKEETLV	SACVQLNMSSPT	PNKDKAVDTPVLADIISITVDLSSFV 1175			
QY	1279	TPQAGEITFSP	PELRIIVATTSSAGSG	SVSSSIPNPKRISAAV	PSGSAATPTTSENKVF 1338			
Db	1176	PEODGTLPLP	PEI-----	IPKGTKLH	SNAI----- 1201			
QY	1339	LTGDLTLIDP	NGEYON--PMLGSD	LDVPLIKLPT-----	NTSDVQVVDLTLSG 1385			
Db	1202	---DLKIIDPT	NVGENHALLSSH	KDIPILSLKTA	EGMTGPTADASLSNIKI--DVSLPS 1257			
QY	1386	DLFPQKGYMT	WTLDSPNQTGLQ	ARWTFDYYRRWVY	IPRONHFEVANSILGSONSMIVVK 1445			
Db	1258	ITPATYGH	TGYWS--ESKME	DGRLLVVGWQ	PTGYK--LNPEKOGALVLLNMLWSHYTDLRAIK 1314			
QY	1446	OGLINNM	LNNARFDDIAYN	FWVSGVTFLAQO	GTPLEEFYSYRGSVAIDAKPRODF 1505			
Db	1315	QEIFAHTIA	ORMELDFSTN	VWVSGLVGVED	QCNIGEDFGFKHLTGVALGLDTQLVEDF 1374			
QY	1506	ILGAAFSKI	VCKTKAIKMHNY	FHKSEYSQASVYGK	FLLNKHOGHALPFIQGV 1565			
Db	1375	LIGGCF	SFOFKTES----	QSYKAKNDVKS	YMGAAIAGIL-----AGPWLKIGA 1419			
QY	1566	VSYGHKHD	TTTLPYSIHER	KNKGWEDLGLA----	DLRISMDLKEPSKDSKRTTVYGE 1621			
Db	1420	FVYGNIN	NDLTDYCTL--G	ISTGWSGCKG	FIAGTSIDYRYIVNPRFISAIYSTVWPVE 1478			
QY	1622	LEYSIROK	QFTEIDYDPRH	FDDCAYRNLSLP	VGCAVEGAIMNCNIIIMYNKLALAYMPSI 1681			
Db	1479	AEYVRIDL	PEISEQCKE	VRTFQKTRFEN	VAIPFGFALEHAYSRGSRAEVNSVQLAYVFDV 1538			
QY	1682	YRNNPV----	CKYRVLSNEAG--	OVICGV--PRT	SARAEYSTOLYLGPFWTLYGNYTIDV 1735			
Db	1539	YRKGPS	VLITUKDAAYS	SWKSYGVDIP	CKAWKARLSNNTWNS--YLSYTLAFNFEWREDL 1596			
QY	1736	GWYTL	SQMTSCGARMIF	1752				
Db	1597	IAYDFNG----	GIRIIF	1609				
RESULT 11								
Q92605	PRELIMINARY; PRT; 1609 AA.							
ID	Q92605							
AC	Q92605;							
DT	01-MAY-1999	(TrEMBLrel. 10, Created)						
DT	01-MAY-1999	(TrEMBLrel. 10, Last sequence update)						
DT	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)						
DE	PUTATIVE OUTER MEMBRANE PROTEIN D FAMILY (POLYMORPHIC MEMBRANE PROTEIN D FAMILY).							
GN	PWP_21	OR CP0897.						
OS	Chlamydia pneumoniae (Chlamydia phila pneumoniae).							
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.							
OX	NCBI_TaxID=83558;							
RN	[1]							









```
Db 1144 MSFVSGDEASAEISNLSVSDQLQHVVYTPDEIEDTY---GHMGDMS-EAKIQDGTLVISW 1199
Qy 1413 TFDYTRRWYIPRDNHFFYANSILGQNSMIVV-----KQGLNNMLNNARFDDIAYN--- 1464
Db 1200 NPTGYR-----LDQKAGALVFNALWEEGAVLSALKNARE---AHNLTA 1240
Qy 1465 -----NFWVSGVGTFF--LAOQGTPLSEERSYYSRGTSVAIDAKPRQDFILGAAPSK 1513
Db 1241 QRMEDYSTNVMGFAFGGFTLSAENLVAIDGYKAGYGGASAGVDIQLMEDFVLGVSGAA 1300
Qy 1514 IVGKTAIKKMHNYFHKGSEVSYQASVYGKFLYLLNKQHGVALPELLOGVVSYGHIKH 1573
Db 1301 FLGKMSQKDAEVSARKG-----VGVSYTGFL-----AGSWFFKGQYSLGETQN 1345
Qy 1574 DTTTLYPSIHERNKGDWEDLGWLADLRIS-MDLKEPSKDS--SKRIIVYGELEYSSIROK 1630
Db 1346 DMKTRYGVLGE-SSASWTSRGVLADALVEYRSLVGPVRPTFYALHFNYPYEVSVASMKEP 1404
Qy 1631 QFTIEDYDPRHFDCCAYRNLSLPGCAVEGAIMNCNLTMYNKLALAYMPSIYRNVPCKY 1690
Db 1405 GFTEQGREARSFEDASLTNTIPLGMKFELAFIKGFQFSEVNSLGSISYAWAEYRKVEGGAV 1464
Qy 1691 RVLSNNSAGVOCVGPRTSARAEYSQLYLGPPTWLYGNVTIDV---GMYT-----LS 1741
Db 1465 QLL---EAGFDWEGAP-MDLPRQELRVALENNTWWSYFTVLGLTAPCGGFTDSTDKLG 1520
Qy 1742 QMTSCGARMIF 1752
Db 1521 YEANTGLRLIF 1531

RESULT 13
Q9RB66 PRELIMINARY; PRT; 930 AA.
AC Q9RB66;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE POLYMORPHIC OUTER MEMBRANE PROTEIN G FAMILY (POLYMORPHIC MEMBRANE
DE PROTEIN G FAMILY).
DE PMP_8 OR CP0307.
GN Chlamydia pneumoniae (Chlamydophila pneumoniae).
OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OC NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,
RA Berry K., Bass S., Linner K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AF002546; BAA98654.1; -;
DR EMBL; AE002193; AAF38164.1; -;
DR TIGR; CP0307; -;
DR InterPro; IPR003357; -;
DR InterPro; IPR003368; -;
DR Pfam; PF02385; OMP; 1.
DR Pfam; PF02415; DUF145; 1.
SQ SEQUENCE 930 AA; 97639 MW; 46A8896761391C09 CRC64;
```

```
Query Match 7.4%; Score 665.5; DB 2; Length 930;
Best Local Similarity 25.0%; Pred. No. 1.7e-27;
Matches 292; Conservative 150; Mismatches 389; Indels 335; Gaps 48;

Qy 658 VSTALSTPSSSTVSSLLTLLAASSOASPATS-----NKETQDPNADTLLIDYVVD 707
Db 10 ISSTLVTP---ILLSIATYGADASLSPDTPDAGGSGTFTPKSTADANG-----TNYVLS 61
Qy 708 TTISKNTAKKGGIYAKKAKMSRIDQLNISENSATEGGGICCKESLELDALVLSVTEN 767
Db 62 GNVYINDAGKG-----TALTG-----FTET 83
Qy 768 LVKREGGGLHAKTVNISNLKSGFSNNKANSSSTGVATTASAPAAAAAASLQAAAAAPS 827
Db 84 -----TGDLLFTGKGYSFSTVDAGSN-----AGAAASTTADKALTFT 122
Qy 828 SPATPTVSGVVGGAIGEYKVFQSC-----SCTCFSGNQAIIDNNPSSSLNVQGA 880
Db 123 GFSNLSTFAAPGTTVASCSTLSSAGALNLTDCNTILFSONVS-----NEANNNGAI 175
Qy 881 YAKTSLSIGSDACTSYIFSGNSYSTGKSQTGOIAGGAIYSPVTLCNCPATFNSNTASI 940
Db 176 TAKT-LSI-----SGNTSSITFTSSAKKLGAIYS-----SAAASISGNTGOL 218
Qy 941 ATPKTSSEDDSGNSIKDTIGGAI---AGTAITLGSVSRFSGNTADLGAAGTTLANATP 997
Db 219 VFNNKGETG-----GGALGFASSTITQNSLSLFTSGNTA-----253
Qy 998 SATSGSONSITEKITLENGSPIFERNOANKRGAIY-----SPSVSIKGN-NITFNQTS 1050
Db 254 -----TDAAGKGAIYCEKGTETPTLTISGNKSLTFAZNS 289
Qy 1051 THDGSAYFTKDATIESLGSVLFTGNVNTATQASSATSGQNTNTANYCAAIFGDPGQTQS 1110
Db 290 VTQGGAI-CAHGLDLSAAGPTLFSNNRC-----GNTAAGKGGAIAIADSGS--- 334
Qy 1111 SQTDAITLLASSGNITFSNNSLONNOGDTPAKFCSTAGYVKL-SLQAAKGTISIFDC 1169
Db 335 -----LSLSANQGDITFLGNTLSTAPTSTRNAIYLGSSAKITNLRAAQGGIYFDP 388
Qy 1170 VHTSKTKGTQNVYETLDINKEENSNP--YGTGTVFSE---LHENK-----SYIQPN 1218
Db 389 IASNT--TGAS---DVLITINQPSNSPLDYSGTIVFSEGEKLSADEAKADNFTSILKQP 442
Qy 1219 AILHNGTILVKEKTELHVVSFEQKEGSKLIMEPCAVILSNQNIANGALAINGLTDLDSMG 1278
Db 443 LALASGTALKGNVELDVNGFTQTEGTLQMOPGTLKAD---TEAISLTKLVVDL:HAL- 498
Qy 1279 TPQAGEIFSPPELRIVATTSASGGSGVSSSIPTNPKRISAAVPSGSAATPTTMEIK-V 1337
Db 499 -----EGNKSYS-----IETAGAIKTI 515
Qy 1338 FLTGDLTLDIPNGNFYQNPMGLSDLDVPLIKLPTNTSDVQVY-DLTLSGDL---FPQKGY 1393
Db 516 TLTSPLVFQDSSGNFYESTHTINQAFQPLVYVFTAATAASDIYIDALLTSPQTPEHYGY 575
Qy 1394 MG-----TWTLDSDNPQTKLARWTFDIY-----RRWYIPRDNHFNAN-----SILG 1436
Db 576 QGHWEATWADTSTAKSGTM--TWVTTGYNPNPERRASVVP--DSLWASFTDRTLQGLMT 631
Qy 1437 SONSMIVVVKGLINNMLNNAFDDIAYNNFWVSCGVTFLAQQGTPLEESYSYRGYSVA 1496
Db 632 SQANSIYQQRGL-----WASGTANFFHKDKSGTNOAFPHKSYGIVG 673
Qy 1497 IDAKPRQDFILGAAFSKIVGKTKAIKKMHNYFHKGSEYSYQASVY--GGKFLYLLAKQH 1554
Db 674 GSAEDFSENIFSVAFCQLFGKDKDLFIVENTSH-----NYLASLYLQHRAFLGLPWPSP 728
Qy 1555 GWA-----LPFLIQGVVSYGHIKHDTTTLYPSITHERNKGDWEDLGWLADLRISMKLEP 1608
Db 729 GSITDMLKDIPLILNAQLSYSYTKNDMDTRVTSYPEA-QGSWTNNSGALELGGSLALYLP 787
```



DE OUTER MEMBRANE PROTEIN 5 PRECURSOR (FRAGMENT).  
GN OMP.  
OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC/CWL-029/VR-1310;  
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;  
RT "Identification of two novel genes encoding outer membrane complex  
associated-surface-layer proteins in Chlamydia pneumoniae.";  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ001311; CAA04671.1; -  
DR InterPro; IPR003357; -  
DR InterPro; IPR003368; -  
DR Pfam; PF02385; OMP; 1.  
DR Pfam; PF02415; DUF145; 1.  
KW Signal.  
FT SIGNAL. 1 15  
FT CHAIN 16 >914  
FT NON\_TER 914 914  
SQ SEQUENCE 914 AA: 95603 MW: 8BF33BAB680FF5E3 CRC64;

Query Match 7.3%; Score 655.5; DB 2; Length 914;  
Best Local Similarity 25.5%; Pred. No. 5.6e-27;  
Matches 282; Conservative 146; Mismatches 433; Indels 247; Gaps 47;

QY 670 VSSLTLAASSQSPATSKNETQDPNADYDLDLDYVVDYTSKNTAKKGGIYAKKAKMS 729  
DB 9 VLSSLTACFTSCSTVFMAENIGPDSFD-----GSTNT-----GYTPKNTT 53

QY 730 RIDQLNISENATEIGGICCKESLELDALVLSLVTENLVGEGGLHAKTVNISMCLKSG 789  
DB 54 GIDYLTGTITLQNLGDSAAALTKGCFSDTTSLSFA-----GKGYLSFLNIKSSAES 106

QY 790 FSES-NKANSSTGVATTASAPAAAASLQAAAAAASP-ATPYSGVGGAIYGERV 847  
DB 107 AALSVTTDKNLSLTGFSS-----LTFLAAPSSVITTPSGKGA----- 144

QY 848 TFSQSGTCQFSGNQAINNPS-----QSSLVQGGAIYAKTSLSGSDAGTSYIFSGNS 903  
DB 145 ---KCGGDLTF-----DNGTILFKQDYCEENGGAISTK-NLSLKNSTGSTS--FEGN- 191

QY 904 VSTGKSQTTGQIAGGAIYPTVTL---NCPATFSNNNTASIAIPTKTSSEDDSGNSIKDTI 960  
DB 192 ---KSSATGKKGKGAICATGTDVNTNTAPTLESNNIAEA----- 228

QY 961 GGAIAGTATILSGVRFSGNTADLGAAGTANANTPSATSGNSITEKITLENGSFIF 1020  
DB 229 GGAINST-----GNCTITGNTS-----LVFSENSVTATAG-----NG----- 260

QY 1021 ERNOANKGAIYSPS-VSTKGN-NITFNONTSHDGSALYFTKDATIES--LGSVLFTCN 1076  
DB 261 -----GALSGDADVTISGNSQVTFSGNQAVANGGAIY-ARKLTASGGGGGTSFSNN 311

QY 1077 NVTATQASSATSGQNTNTANYGAAIFGDPGTQSSQTDAILTILASSGNITFSSNNLSQNN 1136  
DB 312 IVQGT-----TAGNGAI-----SILAAGECSLSAEAGDITFNGNAIVAT 351

QY 1137 QGDTFASKFCSTAGVYKL-SLQAAKGKTSFPDCVHTSKTKTGSTQNVYETLIDNKEE--- 1193  
DB 352 TPQTTKRNSIDIGSTAKITNLRAISCHSIFFDPI-TANTAADST-----DTLNLNKADAG 406

QY 1194 NSNPYTGTVFSESE-LHENKSYIPON-----ALLHNGTLVLKTEKTLHVVSFEQEG 1244  
DB 407 NSTDYSGSIVFSGEKLSEDEAKVADNLSTLTKOPVTLTAGNLVLKRGVTLDTKGFTQTAG 466

QY 1245 SKLIMEPGAVLSNQNTANGALINGLTIDLSMGTPOAGEIFSPPELRIVATTSSASGGS 1304  
DB 467 SSVINDAGTTL---KASTEVEVTLGLSIPVDSLGECK-----KVVIASASAS--- 510

QY 1305 GVSSSIPNPKRISAAVPSGSAATTPTMSKNKVFLLTGLDITLIDPNGNFYQNPMGLGSLDV 1364

Db 511 -----KNVALSGPILLLDNQGNAYENHDLGKTQDF 540  
QY 1365 PLIKL-----PTNTSDYQVYDLTSLGDLFPQKGYMGWTFLD-----SNPQTGKLQARWTF 1414  
Db 541 SFVQLSALGTATTTDPAVPTVATPHY---GYQGTWGTWVDVDTASTPKTKTATLATLT- 596  
QY 1415 DTYRRWYIP---RDNHFYANSILGSONSMIVVKOGLINMLNNAFFDDIAYNNFW/SGV 1471  
Db 597 ---NTGYLPNPEROGPLVPNSLWGSFSDIQAI-QGVIERALTLCSD-----RGFWAAGV 647  
QY 1472 GTFLAQQGTPLSEEFSSYSGTSVAIDAKPRQDFILGAAPSKIVGKTK--AIKKMHYFVH 1529  
Db 648 ANFLDKDKKGEKRYKSGGYAIGGAAQTCSENLSFAFCQLFGSKDFLVAKNHTDTY 707  
QY 1530 KGSEYSYQASVYGKFLYFLLNKOHG-WA-LPFLIOGVVSYGHIKHDTTTTLYPSIHENRK 1587  
Db 708 AGAFYIQTITECSG-FIGLLDLKPGSWSHKPLVLEGGQLAYSHVSNDLKTKYTAYPE-VK 765  
QY 1588 GWMEDLGLADLRISMCLKPEKSKRITVYGELEYSSIRQKQFTTEIDYDPRHFDICAY 1647  
Db 766 GSWGNNAFNMLGASSHSYPEYLHCFDTVAPYIKLNLTVIRODSFSEKGTGGRSFDJNL 825  
QY 1648 RNLSLPVGCAGEGAINMNCNILMYNKLALAYMPSIYRNNPVCKYRVLLSSNEAGOVICVPT 1707  
Db 826 FNLSLPIGVKFE-KFSDCNDFSYD-LTLSYVPLDIRNDPKCTTALVISGASWETAYAINLA 883  
QY 1708 RTSARAEYSTQLYLGPFWTLYGNTIDV 1735  
Db 884 RQALQVRAGSHYAFSPWFVLGQFVFEV 911

Search completed: October 25, 2001, 08:42:40  
Job time: 214 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2001, 08:36:01 ; Search time 28.96 Seconds  
(without alignments)  
4608.350 Million cell updates/sec

Title: US-09-454-684A-180  
Perfect score: 8942  
Sequence: 1 MKWLSATAVFAAALPSVSGF.....IDVGMVTLQMTSCGARMIF 1752

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8827.5	98.7	1751	2 G71518	hypothetical prote
2	6488.5	72.6	1672	2 C81675	polymorphic membra
3	3550	39.7	1770	2 A71517	hypothetical prote
4	3205	35.8	1460	2 D81675	polymorphic membra
5	3112.5	34.8	1723	2 E72067	polymorphic membra
6	3111.5	34.8	1732	2 H86557	polymorphic membra
7	3111.5	34.8	1732	2 C81601	polymorphic membra
8	932.5	10.4	1407	2 B72078	polymorphic outer
9	823.5	9.2	1276	2 B86546	polymorphic outer
10	823.5	9.2	1276	2 C81591	polymorphic membra
11	743.5	8.3	1609	2 A86611	probable outer mem
12	743.5	8.3	1609	2 D72013	polymorphic membra
13	674	7.5	1531	2 H71468	probable outer mem
14	665.5	7.4	930	2 D86546	polymorphic outer
15	665.5	7.4	930	2 A81591	polymorphic membra
16	661.5	7.4	930	2 D72078	polymorphic outer
17	655.5	7.3	928	2 G86546	polymorphic outer
18	655.5	7.3	928	2 C81591	polymorphic membra
19	622.5	7.0	1520	2 A81731	polymorphic membra
20	577	6.5	2232	2 T34434	hypothetical prote
21	555.5	6.2	936	2 C72078	polymorphic outer
22	553.5	6.2	936	2 C86546	polymorphic outer
23	553.5	6.2	936	2 B81591	polymorphic membra
24	530.5	5.9	922	2 B72131	polymorphic outer
25	530.5	5.9	922	2 E86491	polymorphic outer
26	528.5	5.9	922	2 F81539	polymorphic membra
27	517.5	5.8	949	2 F81591	polymorphic outer
28	516.5	5.8	928	2 H86546	polymorphic outer
29	516.5	5.8	928	2 D72077	polymorphic outer

ALIGNMENTS

RESULT 1

G71518  
hypothetical protein pmpB - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: G71518  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
A:Reference number: A71570; MUID:99000809  
A:Accession: G71518  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1751 <ARN>  
A:Cross-references: GB:AE001314; GB:AE001273; NID:g3328833; PIDN:AAC68010.1; PID:g332  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: pmpB

Query Match 98.7%; Score 8827.5; DB 2; Length 1751;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1732; Conservative 8; Mismatches 10; Indels 3; Gaps 2;

Qy	1	MKWLSATAVFAAALPSVSGFCRPEKELNFSRVETSSSTTFTETIGFAGAEYIVSGNASF	60
Db	1	MKWLSATAVFAAALPSVSGFCRPEKELNFSRVETSSSTTFTETIGFAGAEYIVSGNASF	60
Qy	61	TKFTNPTTDTTPTNSNSSSGETASVSDSDSTTTTTPDKGGGAFYNAHSGVLSFMT	120
Db	61	TKFTNPTTDTTPTNSNSSSGETASVSDSDSTTTTTPDKGGGAFYNAHSGVLSFMT	120
Qy	121	RSGETSLSEIKTKEGGAIFSQGELLFTDLTSLTQNNLSQLSGGAIFGGSTLSG	180
Db	121	RSGETSLSEIKTKEGGAIFSQGELLFTDLTSLTQNNLSQLSGGAIFGGSTLSG	180
Qy	181	ITKATFSCNSAEVPAVKKPTEPKAQTASSETSGSSSSGNDVSSPSSRAEPAANLQS	240
Db	181	ITKATFSCNSAEVPAVKKPTEPKAQTASSETSGSSSSGNDVSSPSSRAEPAANLQS	240
Qy	241	HFICATATPAAQTDSTETPSHKPGSGGAIYAKGDLTIADSQEVLFISINKATDKGAIFA	300
Db	241	HFICATATPAAQTDSTETPSHKPGSGGAIYAKGDLTIADSQEVLFISINKATDKGAIFA	300
Qy	301	EKDVSPENITSLKVQTNAGAEKGGAIYAKGDLISQSSKOSLFNSNYSKOGGALYVE3GI	360
Db	301	EKDVSPENITSLKVQTNAGAEKGGAIYAKGDLISQSSKOSLFNSNYSKOGGALYVE3GI	360
Qy	361	NFQDLEIRIKYNKAGTFFTKITLPSLAQASAGNADAWASSPQSGGATTVSDGDS	420
Db	361	NFQDLEIRIKYNKAGTFFTKITLPSLAQASAGNADAWASSPQSGGATTVSDGDS	420

hypothetical prote  
pmp\_3 [imported] -  
polymorphic membra  
probable outer mem  
polymorphic outer  
polymorphic membra  
polymorphic membra  
polymorphic outer  
polymorphic membra  
probable outer mem  
polymorphic membra  
polymorphic outer  
190K surface antig  
polymorphic membra

QY 421 SSGSDTSETVPVAKGGGLYTDKNLSITNITGTHIEIANNKATDVGGAIVKGTGLTCEN 480  
Db 419 SSGSDTSETVPATKAGGGLYTDKNLSITNITGTHIEIANNKATDVGGAIVKGTGLTCEN 478  
QY 481 SHRLQFLKNSDKQGGGIYGEDNITLSNLTGKTLFQENTAKEEGGLFIKGTDKALWTG 540  
Db 479 SHRLQFLKNSDKQGGGIYGEDNITLSNLTGKTLFQENTAKEEGGLFIKGTDKALWTG 538  
QY 541 LDSFCLINNTSEKHGGGAFVTKELISQTVTSDETTPGTPVHGETVITGNKSTGGGGV 600  
Db 539 LDSFCLINNTSEKHGGGAFVTKELISQTVTSDETTPGTPVHGETVITGNKSTGGGGV 598  
QY 601 CTKRLALSNLOSISISGNSAENGGAHTCPDSPPTADTAEPAAASAATSPKSAP-VS 659  
Db 599 CTKRLALSNLOSISISGNSAENGGAHTCPDSPPTADTAEPAAASAATSPESAPVS 658  
QY 660 TALSTPSSSTVSSLLLAASQAASPATSNKETQDPNADTDLIDYVVDTTISKNATKGG 719  
Db 659 TALSTPSSSTVSSLLLAASQAASPATSNKETQDPNADTDLIDYVVDTTISKNATKGG 718  
QY 720 GIYAKKAKMSRIDOLINISENSATEIGGGICCKESLELDALVSLVTENLVGEGGGLHAK 779  
Db 719 GIYAKKAKMSRIDOLINISENSATEIGGGICCKESLELDALVSLVTENLVGEGGGLHAK 778  
QY 780 TVNISNLKSGFSSFNKANSSTGVATTASAPAAAAAALQAAAAAASPATPTYSYVVG 839  
Db 779 TVNISNLKSGFSSFNKANSSTGVATTASAPAAAAAALQAAAAAASPATPTYSYVVG 838  
QY 840 GAIYKEKTVFSCGTCQFSGNQAINDNPSOSSLNVQGGAIYAKTSLSIGSDACTSYIF 899  
Db 839 GAIYKEKTVFSCGTCQFSGNQAINDNPSOSSLNVQGGAIYAKTSLSIGSDACTSYIF 898  
QY 900 SGNVSSTKSGTGTGAIAGAIYSPVTVLNCPATFSNNATASATPKTSSDGGSGNSIKDT 959  
Db 899 SGNVSSTKSGTGTGAIAGAIYSPVTVLNCPATFSNNATASATPKTSSDGGSGNSIKDT 958  
QY 960 IGGAIAGTAITLSGVSREFSGNADLGAAGTLANANTPSATSGQNSITKITLENGSFI 1019  
Db 959 IGGAIAGTAITLSGVSREFSGNADLGAAGTLANANTPSATSGQNSITKITLENGSFI 1018  
QY 1020 FERNOANKRGAIYSPSVSIKGNITFNQNTSTHDSAIYFTKDIATIESLGSVLFITGNVT 1079  
Db 1019 FERNOANKRGAIYSPSVSIKGNITFNQNTSTHDSAIYFTKDIATIESLGSVLFITGNVT 1078  
QY 1080 ATQASSATSGQNTNTANTGAALFGDPGTQSSQTDAILTLASSGNITFSNNSLQNNQGD 1139  
Db 1079 ATQASSATSGQNTNTANTGAALFGDPGTQSSQTDAILTLASSGNITFSNNSLQNNQGD 1138  
QY 1140 TPASKFCSTAGVYKLSLQAAGKKTISFPDCVHTSTPKTGSTQNVYETLDINKEENSPT 1199  
Db 1139 TPASKFCSTAGVYKLSLQAAGKKTISFPDCVHTSTPKTGSTQNVYETLDINKEENSPT 1198  
QY 1200 GTIVFSSSELHENKSIYIPONAILHNGTIVLKEKTELHVYVSPQEKSKLIMFPGAVLSNQ 1259  
Db 1199 GTIVFSSSELHENKSIYIPONAILHNGTIVLKEKTELHVYVSPQEKSKLIMFPGAVLSNQ 1258  
QY 1260 LANGALALNGLTIDUSSMGTPOAGEIFSPPELRIVATTSSASGGGVSSSIPTNPKRTISA 1319  
Db 1259 LANGALVINGLTIDUSSMGTPOAGEIFSPPELRIVATTSSASGGGVSSSIPTNPKRTISA 1318  
QY 1320 AVPSGSAATPTMSENKVELTGLDGLIDPNGNFYONPMLGSDLDVPLIKLPTNTSDVOY 1379  
Db 1319 AAPSGSAATPTMSENKVELTGLDGLIDPNGNFYONPMLGSDLDVPLIKLPTNTSDVOY 1378  
QY 1380 DLTLSGDLFPQGYMGVWTLDNSPQTGKLOARWTFDTPYRWWYIPRONHFFANSLGSON 1439  
Db 1379 DLTLSGDLFPQGYMGVWTLDNSPQTGKLOARWTFDTPYRWWYIPRONHFFANSLGSON 1438  
QY 1440 SMIVVYKQGLINMLNNAFFDDIAYNNFWVSGVGTFLAQQGTPLSEEFYSYRGTSVAIDA 1499  
Db 1439 SMIVVYKQGLINMLNNAFFDDIAYNNFWVSGVGTFLAQQGTPLSEEFYSYRGTSVAIDA 1498  
QY 1500 KPRODFILGAAFSKIYVGTKTKAIKKMHNFYHKGSEYSYQASVYGGKFLYLLNKQHGHALP 1559

Db 1499 KPRODFILGAAFSKIYVGTKTKAIKKMHNFYHKGSEYSYQASVYGGKFLYLLNKQHGHALP 1558  
QY 1560 FLIOGVWSYGHKHDTTTTLYPSIHERNKGWDLGLWDLADLRISMDLKEPSKDSKRITVY 1619  
Db 1559 FLIOGVWSYGHKHDTTTTLYPSIHERNKGWDLGLWDLADLRISMDLKEPSKDSKRITVY 1618  
QY 1620 GELEYSSIRQOFTEIDYDPRHFDCCAYRNLSLPGVCAVEGAIMNCILMYNKLALAYMP 1679  
Db 1619 GELEYSSIRQOFTEIDYDPRHFDCCAYRNLSLPGVCAVEGAIMNCILMYNKLALAYMP 1678  
QY 1680 SIYRNNPVCKYRVLSSNEAGQVIGCVPTRTSARAEYSQLYLGPFWTLYGNVYIDVGMYT 1739  
Db 1679 SIYRNNPVCKYRVLSSNEAGQVIGCVPTRTSARAEYSQLYLGPFWTLYGNVYIDVGMYT 1738  
QY 1740 LSQMTSCGARMIF 1752  
Db 1739 LSQMTSCGARMIF 1751  
RESULT 2  
C81675  
polymorphic membrane protein B/C family TC0694 [imported] - Chlamydia muridarum (stra  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: C81675  
R:Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
A: C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: C81675  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1672 <JET>  
A:Cross-references: GB:AE002338; GB:AE002160; NID:g7190724; PIDN:AAF39510.1; PID:g719  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0694  
Query Match 72.6%; Score 6488.5; DB 2; Length 1672;  
Best Local Similarity 73.0%; Pred. No. 1.9e-301;  
Matches 1292; Conservative 145; Mismatches 214; Indels 119; Gaps 20;  
QY 1 MKWLSATAVPAALVPSVSGFCFPEPKELNFSKVTETSSSTTETIGEGAGAEYIVSGNASF 60  
Db 4 MKWLSATAVPAALVPSVSGFCFPEPKELNFSKVTETSSSTTETIGEGAGAEYIVSGNASF 63  
QY 61 TKFTNI-----PTTD---TTTPTNSNSSSSGGE---TASVSEDSSTTTTTPDKGGGAFY 109  
Db 64 TNFTNIPVKKPTTDDSTSTPTTSSAVDPTKIVPASSSSSPNSGDTSATDPKGGGAFY 123  
QY 110 NAHSGVLSFMTRSCTEGSLTISEIKMTGEGGAIFSQGELLFTDLTSLTQNNLSQSGA 169  
Db 124 NEHSGVLSFMTRSCTEGSLTISEIKMTGEGGAIFSQGELLFTDLTSLTQNNLSQSGG 183  
QY 170 IFGGSTISLSGITATPCNSAEVPAVPVKKPTPEKAQTASETSGSSSSSGNDSVSSPSS 229  
Db 184 IFGGSTISFSGINQATFSNTAEV---VPEETTPNPNGTQTTSQ-----PSPTS- 231  
QY 230 RAEPAAANLQSHFCATATPAATQDTETSTPSHKGSGGAIYAKGDLTIAQSQEVLFSSIN 289  
Db 232 -----KVQSLFYSSSTQANGCADSQTPSHKPGSGGAIYATGDLTISDSQEIFSVN 284  
QY 290 KATKDGGAIFAEDKVSFENITSLKVQTNABEKGGAIAKGDLSIQSSKQSLFNSYSKQ 349  
Db 285 KASKDGAIFAEDKVSFENITSLKVQTNABEKGGAIAKGDLSIQSSKQSLFNSYSKQ 344  
QY 350 GGGALYVEGGINFQDLEERIKYKNAKGTFTTKTLPSELKAQASAGNADANASSPQSGS 409  
Db 345 GGGALYIEGNVDFKDLLEERIKYKNAKGTFTTKTLPSELKAQASAGNADANASSPQSGS 394









Db 61 IVGDITFTFTNIPVPVPTDANDSSSSKGGSSSGATSLIRSNLHSDFDFTKDSYL 120  
QY 87 -----ASVEDSDSTTTTDPKGGGAFY-NAHSG 114  
Db 121 DLYHUFFPSASNTLNPALLSSSSSGSSSSSASAVVAADPKGGAIFYNEANG 180  
QY 115 VLSFTRSGTEGLTLSEIKMTGEGGAIFSGOELLFTDLTSLTIONNLSQLSGGAIF--- 171  
Db 181 TLTFTTDSNPGSLTLQNLKMTGDGAAYISGPLVFTGLKLNLTFTGNESQKSGGAAYTGG 240  
QY 172 -----GG-----STIS 177  
Db 241 ALTQOAVEAVFTGNTSAGOGAIVYKBEATLNFALDSLKFENKTSQAGGGIYTESTIT 300  
QY 178 LSGITKA-TFSCNSAEVPAVPKPTPEKQAOTASEYSGSSSSGNDSSVSPSSRAEPAAA 236  
Db 301 ISNITSIEFINKASVPAPPEPTSP-----APSLINSITIDTSTLTQTRAASATPAVA 355  
QY 237 NLOSHFICATATPAQOTDTETSPSHKPGSGAIIYAKGDLTI----- 278  
Db 356 PV-----AAVTPPTISTQETA-----GNGCAIYAKOGISISTEFKDLTFKSNASVDAT 403  
QY 279 -----ADSQEV-----LFSINKATDGGGAIAEKDVSENTITSLVK 314  
Db 404 LTVDSSTIGESGAIFAADSIQIOCTGTTLFSGNTANKSGGGIYAVGQVTLIEDIANLKM 463  
QY 315 QTNGAEEKGAIYAKGDLDSIOSK-QSLFNSNYKOGGALYVEGGINFDLEEIRIKYN 373  
Db 464 TNNCKEGGAIYTKALTINNGAILTFPSGNTSTONGGAIFAVGGITLSDLVEVRFSKN 523  
QY 374 KAGTETKTRITLPSLKAQASAGNADAWSSPQSGSGATTVSDSGSDSSGSDSTSEYVP 433  
Db 524 KTGNY-----SAPITKA-----ASNTAPVYSSITTAASPAV-----PAAAAAP 561  
QY 434 VT--AKGGGLYTDKNLSITNITGIETIANKATDVGGGAYVKGTLTCENSHRLQFLKNSS 491  
Db 562 VTNAAKGALYSTEGTLVTSKSIILSFENECNOGGGAYVTKTQCSDSHRLOFTSNKA 621  
QY 492 DKOGGIIYGEDNITLSNLTGKTLFOENTAKKEGGGLFKGTDKALTMGLDSPCLNNITS 551  
Db 622 ADEGGGLYCCDDVTLNLGKTLFQENSSKEKGGGLSL-ASGSLTMTSLESFCLNANTYA 680  
QY 552 EKHHGGAYFKISQTYT---SDVETIPGTVPHGTEVITGNKSTGGNGGVCYTKRLALS 608  
Db 681 KENGGGANVPENIVLTFTYTPNEPAPVQVPYGEALVTGNTAT-KSGGGIYTKNAAPS 739  
QY 609 NLOSTISGNSAENGCGAHTCPDSEPTADTAEQAPAAASAASTPKSAPVSTALSTPSS 668  
Db 740 NLSSVTFDQNTSSNGG-----ALLTQKAAD----- 765  
QY 669 TVSSLTLAASQAAPATSNKETQDPNADPTDLLIDVVDPTTISKNTAK-KGGGIYAKKAK 727  
Db 766 -----KTDCSFYITVNITNNTATNGGGIAGGKAH 797  
QY 728 MSRIDQNLISENATBIEGGICCKESLELDALVLSVTENLVKEGGGLHAKTVNISNLK 787  
Db 798 FDRIDNLTVQSNQAKK-GGGVLEDALEKVTIGSVSQTATSGGGIYAKDIOQLALP 856  
QY 788 SGFSFNNKANSSTGCVATTASAPAAAAASLQAAAAAPSSPATPYSGVVGAIYGE-K 846  
Db 857 GSFTITDNKVETSLT-----TSTNLYGGGIIYSSGA 886  
QY 847 VTFSCGSCQFCQSGNOAIDNPNSSQLNVGGAIVAKTSLISGSDAGTSYIFSGNSVST 906  
Db 887 VTLNLSGFTGIGTNSVINTATSQDA-DIQGGGIYATTSLSI--NQCNTPIILFSNNSAAT 943  
QY 907 GKSQTTQIAGGAIYSPVTVL---NCPATFSNNTASIAATPKTSSSDGSGNSIKDTIGGA 963  
Db 944 KKTSTTKIAGGAIFAATAVTIENNSQPIILFNLSAK-----SEATTAATAGN--KDCSCGA 997  
QY 964 IAGTAITLGSVSR--PSGNTADLGAIG--TLANANTPSATSGSQNSITEKITLENGSFI 1019  
Db 998 IAANSVTLTNNPNTFEITKGYAETGAIGCIDLTNGSPPRKVS-----IADNGSVL 1047

QY 1020 FERNOA-NKRGAIVSPSVSIKGNNTIFENONTSTHDGSAIYFTKDATIESLGSVLTGNV 1078  
Db 1048 FODNALNKGGAIVGETIDISRTGATFIGNSKHGSAICCTALTALPNSQIIFENKV 1107  
QY 1079 TATQASSATSGONTANTANYGAAIFGDPGQTQSOTDAILTLASSGNITFSNNSLQNOG 1138  
Db 1108 TETTATTRAS-----INNLAGAIYGNNET-----SDVTISLSAENGSIFFKNLC----- 1152  
QY 1139 DTPASKFCSIAGYVKL-SLOAAKGGKTSFFDCVHSTTKTGSTQNVYETDLINKENSNP 1197  
Db 1153 -TATNKYCSIAGNVFTAEASAGAKAISFYDAVNSTKETNA-----QELKNEKATS-- 1204  
QY 1198 YGTTVFVSELHENKSYIPONAILHNGTLVLKEKTELHVVSFEQKEGSKLIMEPQAVLSN 1257  
Db 1205 -TGTTLFSELHENKSYIPQKVTFAGHNLILGKNELSVSFTQSPGTTITWPGSVLSN 1263  
QY 1258 QNIANGALAINGLTIDLSSMGTPQAGEIFSPPELRIVATSSASGGSGVSSSIPTNPKRI 1317  
Db 1264 HSKEAGGIAINNVIIDFSEIVPTKDNATVAPPTLKVSRTNAD----- 1306  
QY 1318 SAAVSGSNAITPTMSENKVFLTGDLTLIDPNGNFYQNPMLGSLDLDVPLIKLPTWTS-D-V 1376  
Db 1307 -----SKDKIDITGTVTLLDPNGNLYQNSYLGEDRDITLFDNDSASGAV 1351  
QY 1377 QVYDLTSLGDLFPQGYMGTTWLTDSNPQTGKLOARWTFDYRRWVYIPRDNHIFYANSILG 1436  
Db 1352 TATNTVTLQNLGAKKGYLGTWNLDPNSSGSKILKWTFDKYLRWPIPRDNHIFYINSIWG 1411  
QY 1437 SQNSMIVVKQGLINMLNAREDDIAYNNFVSGVGTFLAQOQCTPLSEFYSYRSQTSVA 1496  
Db 1412 AQNSLTVTKQGLILGNLNNAREDPAFNNFWASAIQSFLRKEVSRNSDSFTYHGRGYTAA 1471  
QY 1497 IDAKPRODFILCAAPSKIVGKTKAIKMMHNYPHKSGSEYQASVYCGKFLYELLNKHQHW 1556  
Db 1472 VDAKPROEFILCAAPSQVFEHAESEYHLDNYKHGSGHSTQASLYAGNIIFYPAIKSR-- 1529  
QY 1557 ALPFLIQGVVSYGHKHDPTTLTPSIHERNKGWDELGLADLRISMDLKEPSKSSKRI 1616  
Db 1530 --PILFQGVATYGYWQHDITTYPSIEEKNMANWDSIAWFLDLRFSVDLKEPQHPSTARL 1587  
QY 1617 TVYGELEYSSIRQKQTEIDYDPRHFDCCAYRNLSLPVGCAGEGAIMNCNIMLYNKALAA 1676  
Db 1588 TFYTEAEYTRIRQKFTELDYDPRGSFACSIGNLAIPTGFSVDGALAMREIILYNNKVSAA 1647  
QY 1677 YMPSIYRNPNVCKYRVLSSNEAGQVTCGYPTTSARAEBYSTQIYLGPFWTLXGNTIDVG 1736  
Db 1648 YLPVILRNPNKATYEVLTSTKEKGNVNVNLPTRNAARAESSQIYLGSWTLXGTYTIDAS 1707  
QY 1737 MYTLSQMTSCGARMIF 1752  
Db 1708 MNTLVQMANGIRFVF 1723

## RESULT 6

HB6557  
polymorphic membrane protein B Family [imported] - Chlamydomophila pneumoniae (strain J  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #texC\_change 02-Mar-2001  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: HB6557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1723 <SFO>  
A:Cross-references: GB:BA000008; NID:g8978911; PIDN:BAA98746.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp\_20

Query Match 34.8%; Score 3111.5; DB 2; Length 1723;  
Best Local Similarity 37.6%; Pred. No. 2e-140;  
Matches 727; Conservative 284; Mismatches 528; Indels 397; Gaps 45;

QY 1 MKWLSATAFAAIVLPSVSGFCPEPEKELNFSRVETSSSTT-----FTETIGEA-CAEYI 53  
DB 1 MKWLPATAFAAIVLPSVSGFCPEPEKELNFSRVETSSSTT-----FTETIGEA-CAEYI 60

QY 54 VSNASFTFTNIPDTOTPTTNSNS-----SSSGET-----86  
DB 61 IVGDITFSTFTNIPVPTPDANDSSNSKSGSSSGGATSLRNSNLHSDFTTKDSVL 120

QY 87 -----ASVSEDSOSTTTPDPKGGGAFY-NAHSG 114  
DB 121 DLYHLFPFASNTLNPALLSSSSSGSSSSSSSGSASAVVAADPKGGAAYFNEANG 180

QY 115 VLSFMTFSGTEGLTSELKMTCEGAIFSGOGLLTLTSLTIONNLSGGAIF---171  
DB 181 TLTFTDSDGNPGLTLONLKMGTGDAAIYSGPLVFTGLKLNLTFTGNESOKSGGAAYTEG 240

QY 172 -----GG-----STIS 177  
DB 241 ALTTQAIWEAVFTGNTSAGOGAIVYKATLFLNALDSLKFEKNTSGQAGGGIYTESTLT 300

QY 178 LSGITKA-TFSCNSAEVAPVKKPTPKAQTASETSGSSSSSGNDVSFSSRAEPAA 236  
DB 301 ISNITKSGEIAFISNKASVPAPAPPTSP-----APSLINSTTIDTSTLQTRAASATPAVA 355

QY 237 NLOSHFICATAPPAOTDTETSTPSHKPGSGGAIYAKGLDITI-----278  
DB 356 PV-----AAVTPPTISTOETA-----GNGGAIYAKOGISISTFTKDLTFKSNASVDAT 403

QY 279 -----ADSQEV-----LFSINKATDGGGAFAEKDVSFENITSLKV 314  
DB 404 LTVDSSTIGESGGAIFAADSIQIOCTGTTLFSGNTANKSGGGIYAVGVTTLEDIANLKM 463

QY 315 QTNGABEKGGAIAKGDLSIQSSK-QSLFNSNYSKOGGALYVEGGINFQDLEERIKYN 373  
DB 464 TNNCTCKEGGAIAVTKKALITNGCAILTTESGNTSTDNGGAIFAVGGITLSDLVEVFRSKN 523

QY 374 KAGTFETKTLTSLKAQASAGNADAWASSPQSGCATTVSDSGSSGSDSDTSETVP 433  
DB 524 KTGNY-----SAPITKA-----ASNTAPVWSSSTTAAAPAV-----PAAAAAP 561

QY 434 VT--AKGGGLYTDKNLSITNITGIIIBIANNKATDVGGAIVKGTLCENSHRLOFLKNSS 491  
DB 562 VTNAAGGALYSTEGTLVSGITSLIFENNECONQGGAYVYTFQCSDSHRLQFTSNKA 621

QY 492 DKOGGIYGEDNTLTLNLTGKTLFQNTAKEGGGLFIKGTDKALMTGLDSFCLINNTS 551  
DB 622 ADEGGGLYCGDDVTLNLTGKTLFQNSSEKKGGLSL-ASGKSLTMTSLESFCLNANTA 680

QY 552 EKHGGAFTVKELISQTYT---SDVEITPGITPVHGETVITGNKSTCGNGGGVCTKRLALS 608  
DB 681 KENGGGAPNENIVLTFTTPTTNEPAPVQOQPVYGBALVTGNAT-KSGGGIYTKNAAFS 739

QY 609 NQOSISIGNSAAENGGAHTCPDSPPTADTAEOPAAAAAATSTPKSAPYSTALSTPSSS 668  
DB 740 NLSVTFDQNTSENGG-----ALLTQKAD-----765

QY 669 TVSSLLTLAASQASPATSNKETQDPNADPTDLLIDYVVDPTTISKNTAK-KGGGIYAKKAK 727  
DB 766 -----NTDCSFYITVNTNNTATNGGGGIAGGKAH 797

QY 728 MSRIDQLNSENSATBIGGICCKESLELDALVLSVTENLVGKEGGGLHAKTVNINSLK 787  
DB 798 FDRIDNLTVQSNQAKK-GGGVLTEDALILEKVITGVSQNTATFESGGGIYAKDIQLQALP 856

QY 788 SGFSFNNKANSSTGVAITATAPAAAAAASLQAAAAAASPSPATPYSGVVGGAICYE-K 846  
DB 857 GSFTITDNKVTESLT-----TSTNLVGGGGIYSSGA 886

RESULT 7

C81601

polymorphic membrane protein B/C family CP0212 [imported] - Chlamydocephila pneumoniae  
C:Species: Chlamydocephila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: C81601  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255  
A:Accession: C81601  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1732 <REA>  
A:Cross-references: GB:AR002182; GB:AR002161; NID:g7189140; PIDN:AAF38082.1; PID:g718914  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0212

Query Match 34.8%; Score 3111.5; DB 2; Length 1732;  
Best Local Similarity 37.6%; Pred. No. 2e-140;  
Matches 727; Conservative 284; Mismatches 528; Indels 397; Gaps 45;  
Qy 1 MKWLSATAVFAAVLPSVSGFCFPEPEKELNFRVETSSSTT-----FTETIGEA-GAEYI 53  
Db 10 MKWLPATAVFAAVLPAALTAFGDPASVEISTSHTGSDPTSDAALTGFTQSSSTETDGTYYT 69  
Qy 54 VSGNASFTKFTNIPDTDTTPNSNS-----SSSGCET----- 86  
Db 70 IVGDITFTFTNIPVVPVTPDANDSSNSKGGSSSGATSLIRSSNLHSDPFDFTKDSVL 129  
Qy 87 -----ASVSEDSDSSTTTTDPKGGGAFY-NAHSG 114  
Db 130 DLYHLFFPSASWLNALLSSSSGSSSSSSSSGASAVVAADPKGAIFYNEANG 189  
Qy 115 VLSFMRSGTEGSLPLSEIKWTGEGGAIFSGELLFTDLTSLTIONNLSQSGGAIF--- 171  
Db 190 TLTFTTDSGNPGSLTLQNLKMTGDGAAYISKGPLVFTGLKMLTFTGNSQKSGGAAYTEG 249  
Qy 172 -----ADSOEV-----LFSINKTKDGAIFAEDKVSFENITSLKV 314  
Db 250 ALTTQAIVEAVFTGNTSAGOGGAIYVKEATLFNALDSLKPEKNTSGGAGGIYTESTLT 309  
Qy 178 LSGIITKA-TFSCNSAEVPAVKKPTPEKRAQTASETSGSSSSGNDSDVSSPSSRAEPAAA 236  
Db 310 ISNITKSIEFISNKASVPAPPEPTSP-----APSSLNITIDTSTLQTRAASATPAVA 364  
Qy 237 NLOSHFICATPAPAAQTOTETSTPSHKPGSGGAIYAKGLDIT----- 278  
Db 365 PV-----AAPTPTPISTQETA-----GNGGAIYAKOGISITFEKDLTFKSNSASVDAT 412  
Qy 279 -----ADSOEV-----LFSINKTKDGAIFAEDKVSFENITSLKV 314  
Db 413 LTVDSSTIGESGATFAADSIQIOQCTGTLFSGNTANKSGGGIYAVGOVLTLEIANLKM 472  
Qy 315 QTNGAERGGAIYAKGDLISIOSSK-QSLFNSNYKQGGGAIYVEGGINFQDLEERIKYN 373  
Db 473 TNNTCKEGGAIYTKALTINNGAILTFSGNTSDNGGAIFAVGCIITLSDLVYRFSKN 532  
Qy 374 KAGTFETKITLPSLKAQASAGNADAWASSPQSGGATTVSDSGSSGSDSTSETVP 433  
Db 533 KTGNY-----SAPITKA-----ASNTAPVYSSSTTAASPAV-----PAAAAAP 570  
Qy 434 VT--AKGGGLYTDKLSITNITGIIETIANNKATDVGGAIVKGTITCENSHRELQPLKNS 491  
Db 571 VTNAAGGALYSTEGTLVSGTISILSFENNENQNGGGAIVYTKTQCSDSHRLQFTSKA 630  
Qy 492 DKQGGIYGEDNITSLNITGKTLFOENTAKBEGGLFIKGTDKALTMGTGLDSFCLINNTS 551  
Db 631 ADEGGGLYCGDDVTLNLTKTLFOENSSERKGGGLSL-ASKSLTMTSLESFCLNANTA 689  
Qy 552 EKHGGGAFVTKIEISQTYT---SDVETIPGTPVHGETVITGNKSTGGGCGVCTKRLALS 608  
Db 690 KENGGANVPENIVLTFTYTPTPNEPAPVQQPVYGEALVTGNTAT-KSGGGIYTKNAAFS 748  
Qy 609 NLOQISISGNSAENGGAHTCPDSFPPTADTAEQAPAAASATSTPKSPAVSTALSTPSS 668  
Db 749 NLSSVTFDQNTSSENG-----ALLTQKAD----- 774

Qy 669 TVSSLTLAASSQASPATSNKETQDPNADTOLLIDYVVDVTTISKNTAK-KGGGIYAKKAK 727  
Db 775 -----KDCSFTYITNVNITNTNTATGNGGGIAGGAKH 806  
Qy 728 MSRIDOLNISENSATEIGGGICCKESLELDALIVLSVTENLVGKEGGLHAKTVNISLJK 787  
Db 807 FDRIDNLTVQSNQAKK-GGGVYLEDALILEKVTGVSQNTATESGGGIYAKDIQLOALP 865  
Qy 788 SGFSFNSKANSSTGTVATTASAPAAAAASLQAAAAAPSSPATPTYSGVVGGAIYGE-K 846  
Db 866 GSFITDNKVETSLT-----TSLNLYGGGIYSSGA 895  
Qy 847 VTFSCSCTCOFSGNOAIDNNPSSSLNVOGGAIVAKTSLSIGSSDAGTSYIFSGNSYST 906  
Db 896 VLTNISGTFGTIGNSVINTATSDA-DIQGGIYATVTSLSI--NQCWTPILFNSNSAAT 952  
Qy 907 GKSQTTGQIAGAIYSPVTVL---NCPATFNSNTASIAATPKTSSDSSGNSIKDTIGGA 963  
Db 953 KKTSTTKQIAGGAIFSAAVTIENNSQPIIFLNSAK-----SEATTAATAGN-KDSCGGA 1006  
Qy 964 IAGTAITLGSVSR--FSGNTADLGAAG--TLANANTPSATSGSONSITEKITLNGSFI 1019  
Db 1007 IAAVSITLNNPEITFRKGYAETGAIGCIDLTNGSPPRKVS-----IADNSVL 1056  
Qy 1020 FERNQA-NKRGAIYSPSVSIRKNNITFNQNTSTHDGSAIYFTKDATIESLGSVLFTGNV 1078  
Db 1057 FQDSALNRGAIVGETIDISRTGATFIGNSKHDGSAICCTALTTLAPNSOLIFENKV 1116  
Qy 1079 TATQASATSQNTWANYGAAIFGDPGCTTQSSQTDAILTLASSGNITTFNNSLQNOG 1138  
Db 1117 TETTATTAKS-----INNLAGAIYGNNT-----SDIITSLSAENGSIFFKNNLC----- 1161  
Qy 1139 DTPASKPCSIAGYVKL-SLOAAKGKTISEFDCVHTSTKTKGTSTQNVYETDLDIKENSNP 1197  
Db 1162 -TATNKYCSIAGNKFTAIASAGAKAISFYDAVNVSTETNA-----OELKNEKATS-- 1213  
Qy 1198 YGTIVFSELHENKSYIPQNAIHLNGTLVLKEKTELHVSVFEQKEGSKLIMEPGAVLSN 1257  
Db 1214 -TGTLFSGELHENKSYIPQKVTFAGHNLILGKNAELSVSFTQSPGTTITMGPGSVLSN 1272  
Qy 1258 QNIANGALAIINGLTLDLSSMGTPOAGEIFSPPELRIVATVTTSSAGSGSVSSIPNPKRI 1317  
Db 1273 HSKEAGGTAINNVIIDFSEIVPTKDNATVAPPTLKLVSRTNAD----- 1315  
Qy 1318 SAAVPSGSAATPTPMSENKVFLTGLDITLDPNNGNYQNPMLGSLDVLPLKPTWTSID-V 1376  
Db 1316 -----SKDKIDITGVITLDPNGNLYQNSYLGEDRDITLFLNIDNSASAV 1360  
Qy 1377 QVYDITLSDGLFPQKGYMGWTWLTDSNPQTGLQARWTFDYYRRWYIIPRDNHFYANSILG 1436  
Db 1361 TATNVTLOGNLGAKKYLGTWNLDPNSSSKILKWTDFKYLRFYIPRDNHFYINSIWG 1420  
Qy 1437 SONSMIVYKQGLINMLNARFDDIAYNNFVVGVTFLAQOQGTPLSEFSYISGTSVA 1496  
Db 1421 AQNSLVTYKQGLNMLNARFDDPAFNFWASATGSLRKEVSRNSDSFTYHGRYATAA 1480  
Qy 1497 IDAKPRODFILGAAPSKIVGKTKALKMKNYFHKGEYSYQASVYGGKFLFLNKHQHW 1556  
Db 1481 VDAPRQRFILGAAPSQVFGHAESYHLDNKKHKGSGHSTQSLAGNIFFPAPAIRSR-- 1538  
Qy 1557 ALPFLIQGVSVYGHKHTDITLYPSIHERNKGDWEDGLWDLADLRISMDLKEPSKDSKRI 1616  
Db 1539 --PILFQGVATYGYQHDTTITVYPSIEKNMANWDSIAWFLDLRFSVDLKEPQPHSTARL 1596  
Qy 1617 TVYGELEYSSINOKOFTELDIDPRHFDCAVRNLSLPVGCAGVEGAIMCNILIMYNKLA 1676  
Db 1597 TFYTEAEYTRIOERKFTELDYDPRSFACSACYNLAIPTGFSVDGALAWREIILYNKVSAA 1656  
Qy 1677 YMPSTYRNPKYKVRVLSNNEAGOVICGVPTRTSARAEYSTOLYLGPEWTLGNTYIDVG 1736  
Db 1657 YLPVILRNPKATYEVLTSEKKGNNVNLPTRNAARAEVSSQIYLGSIWTLGTYTIDAS 1716  
Qy 1737 MYTSLQMTSCGARMIF 1752

Db	1717	MNTLVQMANGGIRFVF	1732
<hr/>			
RESULT			
8			
B72078			
Polymorphic outer membrane protein g/i family - Chlamydoiphila pneumoniae (strain CWL029)			
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae			
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000			
C:Accession: B72078			
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Ollinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999			
A>Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.			
A:Reference number: A72000; MUID:99206606			
A:Accession: B72078			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-1407 <ARN>			
A:Cross-references: GB:AF001627; GB:AE001363; NID:g4376721; PIDN:AAD18588.1; PID:g437672			
A:Experimental source: strain CWL029			
C:Genetics:			
A:Gene: pmp_6			
<hr/>			
Query Match 10.4%; Score 932.5; DB 2; Length 1407;			
Best Local Similarity 23.9%; Pred. No. 1.2e+36;			
Matches 443; Conservative 233; Mismatches 627; Indels 549; Gaps 74;			
<hr/>			
QY	1	MKWLL--SATAVFAAFLPSVSGFCPEPKELNFSRVETSSSTFTTETICEAGAAYIVSNGA	58
DB	5	LPWLLTSALLWFSL-----HPLMAANTDLSSDNY-----ENGSS---SGSA	42
QY	59	SFTKFNTIPTDTTPPTNSNSSSSGASVEDSDSTTTTP-DPKGGAGFYNAHSGLVS	117
DB	43	AFTA-----KEISDSAGTYTLTSDVSIINVSALTADKSCFTNT-GGALS	87
QY	118	PMTSGTEGSLTSEIKMTGEGGAIPSGOELL-FYDLSLTION-NLSOLSGAIFGGST	175
DB	88	FV---GADHSELVLQPTIALTHDGAAINTNTALSFSGFSLLLDSAPATGTSGG----	137
QY	176	LSLSGITKATFSCNSAEVPAPVKPTEPKAQTASETSSSSSSGNDSSVSPSSRAEPA	235
DB	138	-----KGATCVNTE-----GGTATFTDNASVTLOKNTSEKDA	171
QY	236	ANLQSHFTCATAPAAQTDTETSPSHKPGSGGAIYAKGDLTI-ADSQBVLFSINKATD	294
DB	172	AVSAYSIDLAKTTAAALLDQNTSK-----NGCALCSANTTVQNSGTVTFSSNTATDK	226
QY	295	GGAIFA-EKVSVFNITSLKYOTNGAEKBGGAIYAKGDLISIOSSKOSLFNSYSKOGGA	353
DB	227	GGGIYSKEKDSTLDANTGVWTFKSNKTAKTGGAWSDDNLALTGNVTQLFOEN-----	278
QY	354	LYVEGINFDLEIRIKNYKAGIFTKKITLPSLKAQASAGNADAASSSQSGSGA---	411
DB	279	-----KTTGSAAQA-----NNPEGCGGATC	298
QY	412	---TTVSDS---GDSSGSDSDTSETPVWTAKGGGLYTDKNLSITNITIGIEIANNKATD	465
DB	299	CYLATADTKTGLAISQNQEMSFSTNT--TTANGGAIYATK-----	336
QY	466	VGGGAYVKGTTTCENSHRLQFLKNSDKO-GGGIYGE-DNIITLSNLTGKTLFOENTAKEE	523
DB	337	-----CTLDGNTTLTFDQNTATACGCCGAIYETEDEFSLKSGTGTVTSTNTAK-T	385
QY	524	GGGLFIKCTDKA----LTMTGLDSFCLINTSEKHGGAGFYRKISQITYTSDVTIPOI	578
DB	386	GGALYSKGNSSLTGNNTLLFGNKRATGPSNSANEGCGGAILAFIDSGSVSDK---TGL	442
QY	579	TPVHGETY-ITGNKSTGGNGVCVKRLALSNLQISISISGNSAAENGGAHGPCDSFEPTA	637
DB	443	SIANNQEVSLTUNAAT-VSGGAIYATKCTLTGNGSLTDFDGTAGTSGGAIYETEDFTL-	500
QY	638	DTAEOPAAAASATSTPKSPAYSTALSTPSSSTVSSLTLLAASSQASPATSNETQDPNAD	697



Db	487	AGGAIYETEDFTLTGSGTVTFSTNTA-----KTGALHTKGN-----TSFT	529
Qy	971	LSGVRSFGSNTADLGAAGITLANANTPSATSGNSI-----TEKITL-ENGSGFI	1019
Db	530	KNKALVFGNSA-----TATATTTTDOECGCGAILCNISEDIATKSLTLTENEELS	581
Qy	1020	PERNQANKR-GAIYSPSVSIKGN-NITFNQNTSTHDGSAIYFTKDATIESGSLVFTGNN	1077
Db	582	FINNTAKRSGGGIYAPKCVISGESINFQGNFAETSGGAIY-SKNLSITANGPVSFTNN-	639
Qy	1078	VTATQASSATSGONTANTYNGAAGIDGCPCTOSSQTDAILTLASSGNITFSNNSLQNNQ	1137
Db	640	-----SGKG-----GAIYADSGE-----LSLEADGDIITFSGN--RATE	673
Qy	1138	GDPSPAKFCISAGYVYKLSLQAARGKTIISPFDCVHTSKTKTGSTQNVYETLDIN-----	1190
Db	674	GTSTPNSIHLGACAKTKTLAARPGHTIYVDITPWFAPASGCT-----IPEIYNVVKVIAV	730

QY	1191	---KEEN-----SRPYGTGVFSS-ELEHNKSPQNAI-----LHNGT	1221
Db	731	PPQPKNGPIASVPVVPVAPANPNTGTVFSSGKLPSQDASIPANTTTILNOKINLAGGN	790
QY	1226	LVLKEXTELHVVSFEQEGSKLIMEGAVL--SNQIANGALAINGLTIDILSSMGTPQAG	1283
Db	791	VLKIGATLQVSYFTQPODSTVFMDAGTTLTTTNTTNDGSDTKLNSVLDALDGRK--	848
QY	1284	EIFSPPELRIVATTSASGGSGVSSSIPTNPKRISAAVPSGSAATPTMTSENKVFLLTGDL	1343
Db	849	-----MTIAVNTSGGLKIS-----GDL	867
QY	1344	TLIDPNGNFYQNMLGSDLDVPLIKLPTNTSDVQVYDLT--LSGDLFPQKGYMGTWTLLDS	1401
Db	868	KFHNNESFYDNPGLKANLNLPLDLSSTSGTVNLDDFNPISSMAAPDYCYQGSWTL--	926
QY	1402	NPQTG-----KLAQRTFTDYR-----RWYIYIPRONHHFYANSILGSONSMIVVQGLINN	1451
Db	926	VPKVGAGGKVTLVAEQALGYTPKPELRATLVP-----NSLWNAYVNIHSIQEIA--	976
QY	1452	MLNNAFFDILAYNNFVWSGVGTFLAQOGTFLGEEFSYSRGTSVAIDAKPRQDFILGAAF	1511
Db	977	---TAMSDAPSHPGIWIIGGIGNAFHQDKOKENAGFRLISRGYIVGGSMTTQEVTFVAVP	1033
QY	1512	SKIVGKTK--AIKKMHNYFHKGS--EYSQVASYVGG---KFLYFLLNKHQHGWAFLDQ	1563
Db	1034	SQLFGSKDYVWSDIKSQVYAGSLCAQSSVYIPLHSLRRHVLSKVLPELPG-ETPLVLH	1092
QY	1564	GVYSYGHKHDTTTLYPSIHERNKGDWEDLGLWADLRISMCLKPEPSKDSRRITVYG---	1620
Db	1093	GOVSYGRHNHNMVTKLAN-NTQCKSDWDSHSPAVEVGGSL-----PVDLUNYRYLTSYSPYV	1147
QY	1621	ELEVSSTRQKOFTEIDYDPRHFDCCAYRNLSLPVGCACV--EGATMNCNLMNYNKLALAYM	1678
Db	1148	KLQVSVNQKGFQGEVAADPRIFDASHLVNVSIPMGLTTFKHESAKPPSALL-----LTGSYA	1203
QY	1679	PSIYRNPNVCYKRVLSNBEAGOVICGVPTRTSARAEYSTQLYLGPPFWLYGNYTIDV3MY	1738
Db	1204	VDAYRDPHC-LTSLTNGTYSWSTFATNLSRQAFEAESCHLKLLHGLDCFASGSCEL3SS	1262
QY	1739	TLSQMTSCGARMIF 1752	
Db	1263	SRSYNANCGRTRYSF 1276	
RESULT	11		
A86611			
probable outer membrane protein D family [imported] - Chlamydomophila pneumoniae			
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae			
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001			
C:Accession: A86611			
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; et al.			
Nucleic Acids Res. 28, 2311-2314, 2000			
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.			
A:Reference number: A86491; MUID:20330349			

RESULT 11  
A86611  
probable outer membrane protein D family [imported] - Chlamydothila pneumoniae  
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: A86611  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; S  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:2030349



A:Accession: A86611  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1609 <STO>  
A:Cross-references: GB:BA000008; NID:g8979337; PIDN:BAA99171.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp\_21

Query Match 8.38; Score 743.5; DB 2; Length 1609;  
Best Local Similarity 22.38; Pred. No. 1.4e-21;  
Matches 406; Conservative 265; Mismatches 687; Indels 459; Gaps 79;

```
Qy 165 LSGAIFGSGTISLS-----GITKATFSCNSA---EVPAPVKKPTEP--KAOTASE----- 210
Db 23 LSAGIAFAHLSHSELDLGVNFVKQFEHSAHVEAQTSLKGSDPVNPSPKSEKVLVT 82
Qy 211 ----TSGSSSSS-----GND-----SVSSPSSSRAEP 233
Db 83 QVPLTQSGSGESLDLADANLEHLEHQLFEETTFVGIDQKLWMSDLDRNFSQPTQ---EP 139
Qy 234 AAANLQSHFICATAPAAQTOTETSPSHKEG-----SGAIYAKGDLTIADSQEVLFSIN 289
Db 140 DTSNAVSEKI--SSDTKENRKOLETEDPSKKSGLKEVSDLPKSPETAFAVAIASEDLISEN 198
Qy 290 KATKD---GGAIF-----AEKDVSEFI-----TSLKVOTNGAEKGGAIYAK 329
Db 199 ISARDPLOGLAFFKNTSSQISSEKSSFGIIFSGGANSGLGFENLKAPKSGAAVYS 258
Qy 330 GDLSTQSSKQSL-FNSNYSKGGALYVEGGINF-----ODLEEIRI-- 370
Db 259 RDIVFENLVKGLSFICESLEDGSA---AGVNIIVTHCGDVTLTDCATGLDLRALRVK 314
Qy 371 KYNKAGTETETK-----ITPLSKAQASAGNADAWASSPQSGSGATTVSDGSDSSG 423
Db 315 DFRSGGAVTARNHEVONNLAGGILSVYVGNKGAIVVEKNSAEKSGAGFACGFSFYISNE 374
Qy 424 SDSDFSEVPVYAKGGGLYTKNLISITNITGIIETANKA-----TD-VGGGAY- 471
Db 375 NYALWKENQALS--GGALISSASDIDIQNCSAIEFSGNOSLIALGEHMLGDFVGGGALA 432
Qy 472 VKGTLTCENSHRLOFLKNSDKOGGIYGEDNITLSNLTGKTLFQENTAKBEGGGLFKG 531
Db 433 AOGTLTRNNAVOCVKNTSKTHGGAILA-GTVDLNETISEVAFKQNTAALTGGAL--SA 489
Qy 532 TDKALTWTGLDSFCLINTSEKHG---GAFVTKSEISQTTSDVETIPGTPVHGETVI 587
Db 490 NDKVLIANNFGEILFEQNEVRNHHGGAIYCGCRSNPKLQKQDSGENINIGNS---GATTF 546
Qy 588 TGNKST-----GGNGGGVCTKRLAL-SNLQSIISGNSAAEN-----GGGAHT 629
Db 547 LKNKASVLEVMTOADYAGGALWGHNVLLDSNSGNIQFIGNIGSTWIGEVYGGGAIL 606
Qy 630 CPDSPTADTA-----EQPAASAAATSPKSAVSTALSTPSSSTVSSLTLLAASSQ 681
Db 607 STDRTYISNNSGDVVFKNGCQCLAQKYVAPQETAPVESDASS-TNKDEKSLNACSHGD 665
Qy 682 ASPATSNKETQDPNADTDLIDYVVDITISKNTAKKGGIYAKKAKMSRIDQLNISENSA 741
Db 666 YPPKTVEEH-----VPPSLLEEHPV---VSSDTIRGGGAILAQ-----HIFITDNTG 709
Qy 742 TEIGGGICCKESLELDALVLSVTENLVGKEGGLHAKTV-NISNLKSGFSFNKANAN-S 799
Db 710 -----NLRFSNGL---GGGEESTVGDALVGGGALLSTNEVNC 746
Qy 800 SSTGVATTASAPAAAAAQAAPSSPATPTYGYVVGGAIXGEKVTFPSCSGTCQFS 859
Db 747 SNQNVFSDNVTNSGCDs-----GGAILAKKY-----DIS 776
Qy 860 GNOAID--NNFSQSLNVQGGAIYAKTSLSTGSSDAGTSYFSGNSVSTGKSQTTGQIAG 917
Db 777 ANHSVEFYVSGSGKF---GGAVCALNE-SVNITD-----NGSAVSFSKNRT--RLGG 822
```

```
Qy 918 GAIYSPTVTLNCPATFSNNTASIAIATPKT---SSBDGSGSGSIKDTIGAGIAIATL--- 971
Db 823 AGVAAPOGSV---TICGNQCNIAFKENFVFGSENQSRGG-----GAIANSVNTQDN 872
Qy 972 -SGVSFRSGNTADLGAAI--GTLANANTPSATGSSQNSITSEKITLENGSFIFERNQANKR 1028
Db 873 AGDILFVSNSTGSYGGAIFVGLV-----ASEGS-NPRTLITIGNSGIDILFAKNSTQTA 925
Qy 1029 GAIYSPSVSIKGNITFNQNTSTHDGSAIYFTKDATIESLGSVLFTGNNVTATQASATS 1088
Db 926 ASL-----SEKDSFGGAIYTONKLIKVNAGNVSYGNR-----APS 962
Qy 1089 GQNTNTANYGAAIFGDPPTTOSSTQDAILTLASSGNITFNSSNLQNNQGTDPASKFCSI 1148
Db 963 GAGVQIA-----DGGT-----VCLEAFGGDILFEGNI--NFDGSFNAIHLGCGN 1003
Qy 1149 AGYKVLQAAKGTISIFFDCVHTSTKTKGTQNYE-----TLDIN---KEENSN 1196
Db 1004 DSKI-VELSAVQDKNIITQDAI---TVEENTIRGLPKDKVSPLSAPSLIFNSKPKQDOSAQ 1059
Qy 1197 PYTGITIVFSSSELHENKSYIPQNAIILHNGTILVLKTEKTELHVVSFEQKESKLIMEPGAVL- 1255
Db 1060 HHECTIRFS---RGVSKIPQIAAQEGTTLALSONAELWLAGLKOETGSSIVLSAGSTLR 1115
Qy 1256 -----SNQNIANGALAIN-----GLTIDLSSMG 1278
Db 1116 IFDSQVDSSAPLPTEENKEETLVSAQVQINMSSPTPNKDKAYDTPVLADIISITVDLSSFV 1175
Qy 1279 TPQAGEIFSPPELRIVATTSSAGSGSVSSSIPTNPKRISAAVSPGSAATPTTSENKVF 1338
Db 1176 PEQDGTLPPLPEI-----IIPKGTKLHNSAI----- 1201
Qy 1339 LTGDLTLIDPNGNYON-PMLGSLDLDVPLIKLPT-----NTSDVOVYDVLTLISG 1385
Db 1202 ---DLKIIDPTNVGYENHALLSSHKDIPLISLKAETAGMTGPTADASLSNIKI-DVSLPS 1257
Qy 1386 DLFPQKVMGTWTLDNSNPOTCKLOARWTFDTPYRRWVYTPRONHFFVANSILGSONSMIVVK 1445
Db 1258 ITPATYGHGTWMS-ESKMEDGRLLVVGWQPTGYK--LNPEKQALVLLNWLWSHYTDLRALK 1314
Qy 1446 QGLINNNLNNARFDIAYNFWVSGVGTFLAQOQGTPLSEEFYSYRSRGTSAIDAKPRODF 1505
Db 1315 QEIFAHTHTIAQRMELDFSTNVWVGSLGVVEDCQNTIGEPDGEFKHLLTYALGLDTQLVDF 1374
Qy 1506 ILGAAFSLIVCKTKAKIMHNYFHKGSEYSYQASVYGKFLYLLNKHQHALPFLIOGV 1565
Db 1375 LIGGCFSOFFCKTES---QSYKAKNDVKSYMGAAYAGIL-----AGPWLKGA 1419
Qy 1566 VSYGHIKHDTTTLPSIHERNKGWEDLGWLA---DLRISMDLKEPSKDSKRTIVYGE 1621
Db 1420 FVYGNINNDLTDYCTL-GISTGSHWKGFIAGTISIDYRYIVNPRFISAIVSTVPVE 1478
Qy 1622 LEYSIRQKQFTEIDYDPRHDDCAYRNLSLPVGCAGEAIMNCNILMYNKLALAYMPSI 1681
Db 1479 AEYVRIDLPEISEQGEKVRTQKTRFENVAIPFGFALSHAYSRGSAEVSQVLAIVFDV 1538
Qy 1682 YRNPV---CKYRVLSSNEAG-QVICGV-PTRTSARAEYSTQLYLGFWTLYGNYYTIDV 1735
Db 1539 YRKGFSVLITLKDAAYSWKSXGVDIPCKAWKARLSNNTENWS--YLSYTLAFNFEWREDL 1596
Qy 1736 GMYTLTQMTSCGARMIF 1752
Db 1597 IAYDFNG-----GIRLIIF 1609
```

RESULT 12

H72013

polymorphic membrane protein D family CP0897 [imported] - Chlamydothila pneumoniae (S  
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: H72013; F81526  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,





QY 1736 GMYTSLQWTSQARMIF 1752  
| : : : : :  
Db 1597 IAYDFNG---GIRIIF 1609

RESULT 13  
H71468  
probable outer membrane protein D - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: H71468  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
A:Reference number: A71570; MUID:99000809  
A:Accession: H71468  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1531 <ARN>  
A:Cross-references: GB:AE001353; GB:AE001273; NID:g3329271; PIDN:AAC68408.1; PID:g332927  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: pmppD

Query Match 7.5%, Score 674; DB 2; Length 1531;  
Best Local Similarity 21.7%; Pred. No. 2.8e-24;  
Matches 419; Conservative 251; Mismatches 659; Indels 502; Gaps 86;

QY 8 AVFAALVPSVSGFC-----PPPEKLNFSRVETSSSTFT 42  
| : : : : :  
Db 17 SWAAILASVGLASCDVLDHAGQSVNELVYVGPQAVLLDQIRDLFVSGKDSQAEGYR 76

QY 43 ETICEAGAEYIVSGNASFTKNTIPTDTTPT-----NSNSSSSGETAS-VSESDSDS 95  
| : : : : :  
Db 77 LIVGDPSPFQEKDADTLPGKVEQSTLFSVTPVVGQVDDQVSSQGLCSFTSSNLD 136

QY 96 TTTTPDPKGGAFYNAHSGVLSFMRSTEGSLTSLSEIKMTGEGGAIFSOCELLFTDLTS 155  
| : : : : :  
Db 137 -----PRDGESFLG----IAFVGDSKAG-ITLTDVKASLSGAALYSTEDLIF----- 179

QY 156 LTIONNLSQLSGGAIFG-STPISLGIKA-----TFCNSAEVPAVKKPTEPKAQTA 210  
| : : : : :  
Db 180 -----EKIKGLEFASCSLEQGGACAAQSIILHDCQGLQV-----KHCTTAVN 223

QY 211 TSGSSSS-----GNDVSSPSSRAEPAANLQSHFICATATPAQTDTEFTPSHRP-- 264  
| : : : : :  
Db 224 AEGSANDHLFGGGAFFVTGSLSGE-----KSLYMPAGDMVNVANCDAISFEGNSANF 277

QY 265 GSGGAIYAKGD-LTIADSQEVLFSINKATKDGGAIFAEKDVSFENITSLKVQTN---GAE 320  
| : : : : :  
Db 278 ANGGAIAASGRVLFVANDKKTSTFIENRAL-SGGAIAASSDIAFONCAELVPKGNCAIGTE 336

QY 321 EK-----GGAIYAKGDLSTQSKQSLFNSNYKQGGALY-----VEGGINFQD---- 364  
| : : : : :  
Db 337 DKGLGGGAISLSGLTLLQGNHGTCDKNESASQGAIFGKNCQISDNQGVVPRDSTAC 396

QY 365 -----LEIRIKYNKAG--TEETKKTITLPSLKAQASAGNADAWASSSPOSGSGATVS 415  
| : : : : :  
Db 397 LGGGAIAAQEIVSQNNQAGLSFEGGK-----ASFGGIACGFSFSGGA----- 441

QY 416 DSGDSSGSDSDTSETPVATKGGGLYTDKNLSITNITGITEIANNKATDVGGAAYKGT 475  
| : : : : :  
Db 442 -----SVLGTIDISKN-----LGAISFSRNL 462

QY 476 LTCENSHRLQFLKNSSDKQGGIYGEDNITLSNLTKTLFOENTAKBEGGLFKGTDKA 535  
| : : : : :  
Db 463 CTTSDLGOMEY-----QGGALFGENISLSNAGVLTIPKDNIVKT-----FASNGKI 509

QY 536 LTMGTGLDCLINNTSEKHGGGAFVTKETISQTYTSDVETIPGTPVHGSETVITGNKSTGG 595  
| : : : : :  
Db 510 L-----GGGAILAT-----GKVEITNNS----- 527

QY 596 NGGVCVTKRLALSNLQISISGNSAAENGGAHTCPDFFPTADTAEQPAAASAATSPK 655  
| : : : : :  
Db 528 -----EGISFTGNARA-----PQALPQOE---EFP-----LFSKKEG 556

QY 656 APYSTALSTPSSSTVSSLTLLAASSQASPATSNKETQDPNADTDL-----IDYVVD 708  
| : : : : :  
Db 557 RPLSSYGSGGAILGREVAIL---HNAAVVEQNRNLCSEEEATLLGCCGGGAVHGMDS 613

QY 709 TISKNTAKKGGGIYAKKAKMSRIDQLNISNSATEIGGICKESLELDALVLSLVTE 768  
| : : : : :  
Db 614 SIYNSVVRGNVNA-----MGCGVSGGALLSKTVOLAGNSVDFSRNI 657

QY 769 VKEGGGLHAKTVNISNLKSGF-SFSNNKANSSSTGVATTASAPAAAASLQAAAAAPS 827  
| : : : : :  
Db 658 ASLGGGALQASEGNCELVDNGYVLFDRNRR----- 688

QY 828 SPATPTYSVVVGAI---YGEKVTFSQCSTCQFSGNOAI----- 864  
| : : : : :  
Db 689 -----VYGGAIISCLRGD-VVISGNKGRVEFKDNIATRLVYEETVEKVEVEPAEQ 738

QY 865 -DNN-----PSOSSLNVQGGAIYAKTSLSIGSSDAGTSYIFSGNSVSTGKSTQTG 916  
| : : : : :  
Db 739 KDNNELSFLGRAEQSFITAAQALFA-----SEGDLS---PESSISEELAKRRECA 788

QY 917 GGAIYSPVTILNCPATFSNNTATPKTSSDSSGNSIKDITIGGAIACTAITLSGVS 976  
| : : : : :  
Db 789 GGAIFAKRVRI-----VDNQEAUVFS-----NNFSDIYGGAI----- 820

QY 977 FSGNTADLGAAGTLANANPSPATSGNSITKILITLNGSFIFERNQANKRGAIYSPV 1036  
| : : : : :  
Db 821 FTG-----SLREEDKLD-----GQI---PEV 838

QY 1037 SIKGN--NITFNQNTS-----THDGSIAVFTKDATI-ESIGSVLFTGNVNTAQASSAT 1087  
| : : : : :  
Db 839 LISGNAGDVFSGNSKRDEHLPHTGGAICTQNLITISQNTGNVLFY-NV-----AC 890

QY 1088 SQQNTNTANGAIIFGDPGTQSSQTDAILTLASSGNITFSNNSLQNNQGDTPASKFCS 1147  
| : : : : :  
Db 891 SG-----GAVRIEDHGN-----VLLEAFGDDIVFKGNSFRAQG-SDAIYFAG 932

QY 1148 IAGVVKLSLQAAGKKTISFFDCV---HTSKTKTGSTQNVYETLDINKEENSNPYTGIVF 1204  
| : : : : :  
Db 933 KESHI-TALNATEGAHIVFDHALVFENLEERKSA-----EVLINRENPY-YTGSIRF 984

QY 1205 SSELHENKSIYIPQAIILHNGTILVLEKTELHVVSFEQEGKSLIMEPGAVLSNQNIANGA 1264  
| : : : : :  
Db 985 L-----EAESKVPQCIHVQOQSLELLNGATLCSYCFKQDAGAKVLAAGAKLK----- 1032

QY 1265 LAINGLTIDLSMGTGP-OAGEIFSPPELRIVATTSSAGGS-----GVSSSIP-TNPKRI 1317  
| : : : : :  
Db 1033 -----ILDSCTPVQOQHAISKPEAETESSEPEGAHSLWIAKNAQTTPVPMVDIHTI 1083

QY 1318 SAAVPS-GSAATTPMTSENKVFLTG-----DLTIDPNGNFYON-PMLGSDLDVPL 1366  
| : : : : :  
Db 1084 SVDLASFSQOQETVEAPQIVPGSVYRSSELNLELVNTGTGYNHALLKNEAKVPL 1143

QY 1367 IKLPT-----NTSDVQVYDLT--LSGDLFPQKGYMGWTWILDSNPQTKLQARW 1412  
| : : : : :  
Db 1144 MSFVASGDEASAEISLNSVSDLOIHVVTPEIEEDTY---GHMGDWS-EAKIQDGTLVISW 1199

QY 1413 TFDYTRRWYIPRDNHIFYANSILGSONSMIVV-----KQGLINNNLNARPDITAYN--- 1464  
| : : : : :  
Db 1200 NPTGYR-----LDPQKAGALVFNALMEGAVLSALKNARF---AHNLTA 1240

QY 1465 -----NFWVSGVGTFF--LAQGGTPLSEFSYSRGTSVAIDAKPRQDFILGAAFSK 1513  
| : : : : :  
Db 1241 QRMEDYSTNVWGFAFGGFRFLTSAENLVAIDGYKAGVGASAGVDIQLMEDFVLGVSAA 1300

QY 1514 IVGTIKAIKMHNYFHKSGSEYYSQASVYGGKFLYFLNKHQHWALPFLIQWWSYGHK 1573  
| : : : : :  
Db 1301 FLGRMDSQKFDPAEYSRKG-----VVGSVYTGFL-----AGSNFFKQOYSLGETON 1345

QY 1574 DTTTLTYSIHERNKGDWEDLGNLADLRIS-MDLKEPSKDS--SKRITVYGELEYSSIRQK 1630

Db 1346 DMKTRYGVIGE--SSASWTSRGLADALVYRSLVGVVRPTFYALHFNPNPYEVSASMKEP 1404  
QY 1631 QFTEIDYDPRHFDCCAYRNLSPVGCAGVAGTMCNLMYKALAYMPSIYRNNPVCKY 1690  
Db 1405 GTEQGRARSFEDASLTNITPLGKMFELAFIKGOFSEVNSLGISYAWEARKKVEGGAV 1464  
QY 1691 RVLSSNEAGQVIGVPTRTSARAEXSTQLYLGFPPFTLYGNYTIDV----GMYT-----LS 1741  
Db 1465 QLL---EAGFDWEGAP-MDLPRQLRVALENTEWSSYFSTVLGLTAFCGGFTSTDSKLG 1520  
QY 1742 QMTSCGARMIF 1752  
Db 1521 YEANTGLRLIF 1531

RESULT 14  
D86546  
polymorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strain  
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: D86546  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: D86546  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-930 <STO>  
A:Cross-references: GB:BA000008; NID:g9978818; PIDN:BAA98654.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp\_8

Query Match 7.4%; Score 665.5; DB 2; Length 930;  
Best Local Similarity 25.0%; Pred. No. 3.4e-24;  
Matches 292; Conservative 150; Mismatches 389; Indels 335; Gaps 48;

QY 658 VSTALSTPSSSTVSSITLLAASSQASPATS-----NKETQDPNADTLLIDYVVD 707  
Db 10 ISSTLVTP---ILLSTATYGADASLSPDSDFCAGGSGTFTPKSTADANG-----TNYVLS 61  
QY 708 TTISKTAKKGGIYAKKAKMRIDQLNLSSENSATEIGGICCKESLELDALVLSLVTEN 767  
Db 62 GNVYINDACKG-----TALTG--CC-----FTET 83  
QY 768 LVGKEGGLHAKTVNLSLKSFSFSSNKKANSSTGCVATTASAPAAAAAALQAAAAAPS 827  
Db 84 -----TGDLTFTGKGYSFNFVDAGSN-----AGAAAATTADKALTFT 122  
QY 828 SPATPYSGVGVGAIYGERKVTFSOC-----SGTCFSGNOAIDNPNQSSSLNVQGGAI 880  
Db 123 GFSNLSFIAAPGTVVASGKSTUSSAGALNTDNGTILFSONVS-----NEANNNGGAI 175  
QY 881 YAKTSLIGSDAGTSYIFSGNSVSTGKSQTTQIAGGAIYSPVTVLNCPATFSNNNTASI 940  
Db 176 TAKT-LSI-----SGNTSSITFTSNSAKKLGAIYS-----SAAASISGNTGQL 218  
QY 941 ATPKTSSEDSGSGNSIKDTIGGAI---ACTAITLGSVRSFSGNTADLGAAGTLANANTP 997  
Db 219 VPMNNKGETG-----GGALGFASSTITQNSLFFSGNTA----- 253  
QY 998 SATSGSQNSITEKITLENGCSFIFERNQANKRGIY-----SPSVSIKGN-NITFNQNTS 1050  
Db 254 -----TDAAGKGAIYCEKTGETPTTLTISGNKSLTFAENS 289  
QY 1051 THDGSIAIYTKDATIESLGSVLFTGNVTATQASSATSONTNTANYGAAIFGDPGTTQS 1110  
Db 290 VTQGGAI-CAHGLDLSAAGPTLFSNNRC-----GNTAAGKGGAIATDSGS--- 334  
QY 1111 SOTDAILTLASSGNTTFSNNSLQNNQGDTPASKFCFSIAGYVKL-SLOAAKGKTTISFFDC 1169

Db 335 -----LSLSANGDITFLGNTLTSTASPTSTRNAYILGSSAKITNLRAAQGOSIYIDP 388  
QY 1170 VHTSTKTKSTQNVYETLIDINKENSNP--YGTGTIVFSSE---LHENK-----SVLPQN 1218  
Db 389 IASNT--TGAS-----DVLINOPDSNPLDYSCTIIVFSFGEKLSADAKAADNTSILKOP 442  
QY 1219 AILHNGTILVLKKEKTELHVVSFEQEGSKLIMEPAGVLSNQNIANGALAINGLTIDLSMGM 1278  
Db 443 LALASGTALKGNVELDVNGFTQTEGSTLLMQPGTKLAD---TEAISLTKLVVDLSAL- 498  
QY 1279 TPOAGEIFPPELRIVATTSSASGGSGVSSSIPTNPKRISAAVPSGSAATPTPMSENK-V 1337  
Db 499 -----EGNKSVS-----IETAGANKTI 515  
QY 1338 FLTGDLTLIDPNGNFYQNPMLGSDLDVPLIKLPTNTSDVOVY--DLTSLGDL---FPCKGY 1393  
Db 516 TTTSPVLVFDQSSGNFYESTHTINQAFQPLVFTAATAASDIYIDALLTSPVQTPPEPHYGY 575  
QY 1394 MG-----TWTLDSPNQTKLQARWTFDY-----RRVVIYPRDNHFNAN-----SILG 1436  
Db 576 QGHWEATWADTSTAKSGTM--TWVTTGYNPNPERRASVVP--DSLWASFTDITLQGIWT 631  
QY 1437 SQNSMTLVVQGLINNMNARNPDDIAYNFWVSGVGTFLAQOQTPLPSEEFYSYRGTISVA 1496  
Db 632 SQANSIYQORGL-----WASGTANFFHKDKSGTNOAFRHKHSYGIYV 673  
QY 1497 IDAKPRQDFILGAAFSKIYVTKTKAIKKMHNYPHKSGSEYVQASVY--GCKFLYFLLNKQH 1554  
Db 674 GSAEDFSENIQFVAFQCLPGKDKDLFIVENTSH-----NYLASLYLQHRAFLGGLPVPSP 728  
QY 1555 GWA-----LPFLIQGVWSYGHIKHDTTLYPSIHERNKGDWEDGLWADLRISMDLKEP 1608  
Db 729 GSITDMKDIPLILNAQLSYSTYTKNDMDRTYTSYPEA-QGSWTNNSGALELGGSLALYLP 787  
QY 1609 SKDSSKRITVYGELEYSSI--ROKQFTEIDYDPRHFDCCAYRNLSPVGCAGVAGTANCN 1666  
Db 788 -KEAPFGQGYFPFLKQAVYVSRQNEKESGAERAFDDGDLVNCSPVIGIRLE-----K 840  
QY 1667 ILMYNK-----LALAYMPSIYRNNPVCKYRVLSSNEAGQVIGCVPTRTSARAEXSTOLYLG 1722  
Db 841 ISEDEKNNFELISLAYIGDVYRKNRPSRTSLMVSGASWTSCLKNLARQAFSLASAGSHLTLS 900  
QY 1723 PWTLYGN-----YTIDVGM 1737  
Db 901 PHVELSGEAYELRGAHIYNYDCGL 926

RESULT 15  
A81591  
polymorphic membrane protein G family Cp0307 [imported] - Chlamydomphila pneumoniae (s  
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: A81591  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: A81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-930 <REA>  
A:Cross-references: GB:AE002193; GB:AE002161; NID:g7189234; PIDN:AAF38164.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: Cp0307

Query Match 7.4%; Score 665.5; DB 2; Length 930;  
Best Local Similarity 25.0%; Pred. No. 3.4e-24;  
Matches 292; Conservative 150; Mismatches 389; Indels 335; Gaps 48;

```
Qy 658 VSTALSTPSSSTVSSLLTLLAASSQASPATS-----NKETODPNADTDLIDYVD 707
Db 10 ISSTLVTP---ILLSIATYGADASLSPDSDGAGGTFTPKSTADANG-----TNYVLS 61
Qy 708 TTISKNTAKGGIYAKAKMSRIDQLNISENSATEIGGGICCKESLELDALVLSVTEN 767
Db 62 GNVIINDAGK-----TALTG---CC-----FTET 83
Qy 768 LVKREGGLHAKTVNISNLKSGFSNNKANSSTGVATTASAPAAAAAASLQAAAAAPS 827
Db 84 -----TGDLFTFGKGYSEFNTVDAGN-----AGAASTTADKALTFT 122
Qy 828 SPATPTYGVVGGAIYGEKVTFSQC-----SGTCQFSGNQAIIDNNPSQSLNVQGGAI 880
Db 123 GFSNLSFIAAPGTTVASCKSTLSSAGALNLTGNTILFSONVS-----NEANNNGGAI 175
Qy 881 YAKTSLSIGSSDAGTSYIFSGNSYSTGKSQTGGIAGAIYPTVTLNCPATFSNNTASI 940
Db 176 TAKT-LSI-----SGNTSSITFTSNSAKKLGAIYS-----SAAASISGNTGQL 218
Qy 941 ATPKTSSEDSGSGNSIKDTIGGAI---AGTAITLGSVSRKSGNTADLGAAGITLANANTP 997
Db 219 VFMNKGEGT-----GGALGFEASSITQNSLFFSGNTA----- 253
Qy 998 SATSGSONSITEKITLENGSFIFERNQANKRGAIY-----SPSVSIKGN-NITFNQNTS 1050
Db 254 -----TDAACKGGAICYEKTGETPTLTISGNKSLTFAENSS 289
Qy 1051 THDGSAYFTKDATIESLGSVLFTGNVNTATQASSATSGQNTANTYGAALFGDPGTQS 1110
Db 290 VTQGGAI-CAHGLDLSAAGPTLFSNNRC-----GNTAACKGGAIAIADSGS--- 334
Qy 1111 SOTDAILTLASSGNITFSNNSONNOGDTPASKFCSTAGVVKL-SLQAAKGKTIISPPDC 1169
Db 335 -----LSLSANQGDITPLGNTLTSTAPTSTRNAILGSSAKITNLRAOQSIYFYDP 388
Qy 1170 VHTSTKKTGSTQNVYETLIDINKEENSP--YTGTVFSE---LHENK-----SYIPQN 1218
Db 389 IASNT--TGAS---DVLTIINOPDSNPLDYSGTIVFSGEKLSADEAKAADNFTSILKQP 442
Qy 1219 AILHNGTILVLEKTELHVVSFEQKEGSKLIMPGCAVLNSQNIANGALAINGITDLSMG 1278
Db 443 LALASGTLALKGNVELDVGFTQEGSTLLMOPGTLKAD--TEATSLTKLVVDLSAL- 498
Qy 1279 TPQAGEIFSPPELRIVATTSSAGSGSVSSIPTNPKRIASAAPSGSAATPTMSNK-V 1337
Db 499 -----EGNKSVS-----IETAGANKTI 515
Qy 1338 FLTGDLTLIDPNGFYQNPMLGSLDVLPLIKLPTNTSDVQVY-DLTLSGDL---FPQKGY 1393
Db 516 TLTSPLVFQDSSGNFYESHNTINQAFQPLVVVTAATAASDIVIDALLTSPVOTPEPHYGY 575
Qy 1394 MG---TWTLDSNPQTGLQARWTFDTY-----RWVYIPRDNHFYAN-----SILG 1436
Db 576 QGHWEATWADTSAKSGTM--TWVTTGYNPNPERRASVVP--DSLWASFTDRTLQOIMT 631
Qy 1437 SQNSMIVVKQGLINMLNNAARFDDIAYNFWVSVGVTFLAQOQTPLSEEFYSYRGSVA 1496
Db 632 SQANSIYQOGL-----WASGTANEFHKDKSGTNOAFPHKSYGYIVG 673
Qy 1497 IDAKPRODFILGAAFSKIVGKTAKIKMHNYFHKGEYSYQASVY--GCKFLYFLLNKQH 1554
Db 674 GSAEDFSENIFSVAPQLFGKDKOLFIVENTSH-----NYLASLYLQHRAFLGGLPMPSPF 728
Qy 1555 GWA-----LPFLIOGVVSYGHKHDTTILYPSIHERNKGDWEDLGLWADLRISMDLKEP 1608
Db 729 GSITDMLKDIPLILNAQLUSYTYTNDMDTRYTSYPEA--OGSWTNNSGALELGGSLALYP 787
Qy 1609 SKDSKRTIVYGELEYSSI--RQKQTEIDYDPRHFDCAVYRNLNLSLPVGCAGEAIMNCN 1666
Db 788 -KEAPFFQGYFFFLKFOAVYSRQONFESGAFAFAFDGDLVNCISIPVIRLE-----K 840
Qy 1667 ILMYNK-----LALAYMPSIYRNNPVCKYRVLSSNAGQVIGVPTRTSARAERYSTQLYLG 1722
```

```
Db 841 ISEDEKNFELSLAYIGDVYRKNPRESRTSLMVSGASWTSLCKNLARQAFLASAGSHLTS 900
Qy 1723 PFWTLYGN-----YTIDYGM 1737
Db 901 PHVELSGEAAYELRGSAAHIYNVDCGL 926
```

Search completed: October 25, 2001, 08:41:50  
Job time: 349 sec

EMBL:AE001361

ID AE001361 standard; DNA; PRO; 5601 BP.  
XX  
AC AE001361; AE001273;  
XX  
SV AE001361.1  
XX  
DT 22-JUL-1998 (Rel. 56, Created)  
DT 01-NOV-2000 (Rel. 65, Last updated, Version 5)  
XX  
DE Chlamydia trachomatis section 1 of 87 of the complete genome.  
XX  
KW .  
XX  
OS Chlamydia trachomatis  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
XX  
RN [1]  
RP 1-5601  
RX MEDLINE; 99000809.  
RX PUBMED; 9784136.  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia  
RT trachomatis";  
RL Science 282(5389):754-759 (1998).  
XX  
RN [2]  
RP 1-5601  
RX MEDLINE; 99206606.  
RX PUBMED; 10192388.  
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";  
RL Nat. Genet. 21(4):385-389 (1999).  
XX  
RN [3]  
RP 1-5601  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;  
RT ;  
RL Submitted (20-MAY-1998) to the EMBL/GenBank/DBJ databases.  
RL Program in Infectious Diseases, University of California, 235 Warren Hall,  
RL Berkeley, CA 94720-7360, USA  
XX  
DR SPTREMBL; 084881; 084881.  
DR SPTREMBL; 084882; 084882.  
DR SPTREMBL; 084883; 084883.  
XX  
FH Key Location/Qualifiers  
FH  
FT source 1..5601  
FT /db\_xref="taxon:813"  
FT /organism="Chlamydia trachomatis"  
FT /strain="D/UW-3/CX"  
FT CDS 323..640  
FT /codon\_start=1  
FT /db\_xref="SPTREMBL:084881"  
FT /transl\_table=11  
FT /gene="CT873"  
FT /product="hypothetical protein"  
FT /protein\_id="AAC68471.1"  
FT /translation="MSSNRMLATIKKITVLLLSKRKAGIRIDYCALALDAVEYSYLWFL  
FT VVGYAVYSSGVRWEAILEGETLSPGDRAQLDASLESAQVRLAGLMLDYWDGDSRLECKK  
FT I"  
FT CDS complement(815..3451)

```

FT      /codon_start=1
FT      /db_xref="SPTREMBL:O84882"
FT      /transl_table=11
FT      /gene="pmpI"
FT      /product="Putative Outer Membrane Protein I"
FT      /protein_id="AAC68472.1"
FT      /translation="MRPDHMFCCCLCAAILSSSTAVLFGQDPLGETALLTKNPNHVVCTF
FT      FEDCTMESLFPALCAHASQDDPLYVLGNSYCWFSKLHITDPKEALFKEKGLSIQNFR
FT      FLSFTDCSSKESSPSIIHQKNGQLSLRNGSMSFCRNHAEGSGGAI SADAFSLQHNLYLF
FT      TAFEENSSKGNNGGAIQAQTFSLSRNVSPISFARNRADLNGGAICCSNLICSGNVNPLFF
FT      TGN SATNGGAICCSIDLNTSEKGSLSLACNQETLFASNSAKEKGGAIYAKHMLVLRNGP
FT      VSFINNSAKIGGAIAIQSGGSL SILAGEGSLVLFQNNQSRTSDQGLVRNAIYLEKDAILS
FT      SLEARNGDILFFDPIVQESSKESPLPSSLOASVTSPTPATASPLVIQTSANRSVIFSS
FT      ERLSEEEKTPDNLTSQLOQP IELKSGRLVLKDRVLSAPSLSDQPQALLIMEAGTSLKT
FT      SSDLKLATLSIPLHSLDTEKSVTIHAPNLSIQKIFLSNSGDENFYVELLSKEQNNIP
FT      LLTLSKEQSHLHLPDGNLSSHFGYQGDWTF SWKDSDEGHSLIANWTPKNYVPHPERQST
FT      LVANTLWNTYSDMQAVQSMINTIAHGGAYLFGTWGSAVSNLFYAHDSGKPIDNWHHRS
FT      LGYLFGI STHSLDDHSFCLAGQLLGKSSDSFITSTETTSYIATVQAQLATPLMKISAQ
FT      ACYNESIHELKTKYRSFSKEGFGSWHSAVSGEVCASIPVSNNGSGLFSSFSIFS KLQG
FT      FSGTQDGFEESSGEIRSFSA SFRNISLPMGITFEKKSQKTRNYYYFLGAYIQDLKRDV
FT      ESGPVVLLKNAVSWDAPMANLDSRAYMFRLTNQRALHRLQTLLNVSYVLRGQSHSYS LD
FT      LGTTYRF"
FT      CDS      3725..5500
FT      /codon_start=1
FT      /db_xref="SPTREMBL:O84883"
FT      /transl_table=11
FT      /gene="CT875"
FT      /product="hypothetical protein"
FT      /protein_id="AAC68473.1"
FT      /translation="MSIRGVGGNGNSRIPSHNGDGSNRRSQNTKGNNKVEDRVCSLYSS
FT      RSNENRESPYAVVDVSSMIESTPTSGETTRASRGVFSRFQRLVRVADKVRRAVQCAWS
FT      SVSTRSSATRAAESGSSRTARGASSGYREYSPSAARGLRMFTDFWRTRVLRQTS PM
FT      AGVFGNLDVNEARLMAAYTSECADHLEANKLAGPDGVAAAREIAKRWEQVRDLQDKGA
FT      ARKLLNDPLGRRTPNYQSKNPGEYTVGNMFMFYDGPQVANLQNVDTGFWLDMSNLSDVVL
FT      SREIQTGLRARATLEESMPMLENLEERFRRLQETCDAARTEIEESGWTRRESASRMEGDE
FT      AQGPSRAQQAQSFVNECNSIEFSFGSFGEHVRVLCARVSRGLAAAGEAIRRCFSCKKG
FT      STHRYAPRDDLSPEGASLAETLARFADDMGIERGADGTYDIPLVDDWRRGVPSIEGEGS
FT      DSIYEIMMPIYEVMDMDLETRRSFAVQQGHYQDPRASDYDLPRASDYDLPRSPYPTPPL
FT      PPRYQLQNMDVEAGFREAVYASFVAGMNYVVTQPQERIPNSQQVEGILRDMLTNGSQT
FT      FRDLMRRWNREVDRE"
XX
SQ

```

Sequence 5601 BP; 1664 A; 1030 C; 1382 G; 1525 T; 0 other;

tctaggtgag	ctcgtgaagaa	gacgtattgt	aagaacagtt	tctataattt	ggtggtaatt	60
tttattttta	ataataaaaa	atacgcccta	attcagctaa	aacagccttg	tttttattaa	120
acccattttg	ttttctaacg	gaatcgaggt	gaatcatggt	ttgtccttta	cgcagttgct	180
gtactactca	taccagcatg	gagagcgtgt	ctttggaaga	gacctcgctt	cttttagaag	240
cgatatcaca	ggctgtgcaa	acatgcttaa	ccgtagccaa	gcacacagag	aacgctaaga	300
tttataatca	aaaacaaaaga	agatgtcatc	aaacaggatg	ctggcaacaa	ttaaaaaaat	360
tactgtgttg	ctgttgtcca	aaagaaaagc	tggtatcagg	atcgattact	gcgctctcgc	420
tctggatgca	gtcgaatatt	cctacttatg	gttttttagt	gtaggttatg	cggatatatg	480
ttcaggagta	cgttgggaag	cgatcttgga	aggagagact	ttatctcctg	gagatagagc	540
gcaactagat	gcctcttttag	agagtgtcca	agttaggtta	gctggtttga	tgttggacta	600
ttgggacgga	gattctcgcc	tagaatgcaa	gaagatataa	tcgccagggt	tcgagacaaa	660
gtcgggatct	gtgagcttac	agatgcggtt	gctctgaaac	acttctgtcg	gagaacttaa	720
atgtgatacg	cgacgtgaga	aaaaaatgaa	aagtcttaga	gagagatccc	taagactttt	780
ctcaagctct	aaaacgcaat	gctaggcaag	ggaactagaa	cctgtaagtg	gtccccagat	840
ccagggagta	actatgctt	tgcccgcgca	gtacgtaaga	cacatttaac	agcgtctgaa	900
gtctatgcag	agctctttga	ttcgtaagcc	tgaacatgta	ggctcgcgaa	tccaagttcg	960
ccataggagc	atcccaggag	acggcatttt	tgagtaacac	tacaggtecc	gattccacat	1020
cacgttttag	gtcttggtatg	taggctccca	gaaagtaata	gtagtttcgt	gttttttggg	1080
attttttttc	aaatgttatt	cccataggaa	gtgaaatatt	tctgaaagag	ctggcagaaa	1140
aggaccgaat	ctctcccga	ctctcctcaa	aaccgtcctg	tgttcctgaa	aatccttgca	1200
gttttagagaa	aatagagaag	gagctgaaca	gtccggaacc	attggataca	ataggaatcg	1260
atgcgcacac	ttctccgat	actgcaacgc	tatgccagga	tccgaatcct	tcttttagaga	1320
aggagcgata	ttttgttttt	agctcatgga	tactttcatt	atagcatgcc	tgtgcagaga	1380
ttttcattag	aggggtagcg	agttgcgctt	gtacagtagc	tatataggag	gtcgtttctg	1440
tagacgtaat	aaaggaatcg	gacgatttcc	cgagtaattg	tcctgcagcc	aagcagaaa	1500

aatggtcatc	taaactgtga	gtactgatac	cgaataggta	gccaaggcct	ctatgatgcc	1560
aattatcgat	agggttccca	gagctgtcgt	gagcatagaa	taaattagaa	acagcagatc	1620
cccacgttcc	aaatagatag	gctcctccgt	gcgctattgt	attaatcabc	gactgcacag	1680
cttgcataac	ggaatagggtg	ttccaaagag	tggtcgcaac	gagtgtagat	tgacgtttctg	1740
gatgaggcac	atagttttta	ggcggtccaa	tagcaatcag	agaatgccct	tcatacagaat	1800
ctttccaaga	aaaagtccaa	tctccttgat	atccaaagtg	agaagagagg	ttcccatcag	1860
gaagatgtaa	atgagattgc	tcttttagaga	gagtaaggag	aggaatattg	ttttgctctt	1920
tactgagaag	ctctacattt	tcataaaaaat	tctcatctcc	agaattagag	aggaagatct	1980
tttgataga	aagggttaggg	gcgtggatag	ttacgctttt	ttcagtatct	aaggaatgaa	2040
ggggaatact	tagcgtagct	aacttcaaat	cagaggaagt	ttttaaagaa	gttcccgcct	2100
ccataatgag	gagagcttga	ggatccctgag	agagagaagg	cgcggaagg	acagcgcgat	2160
cttttaaaac	taagcgctcg	gatttccagtt	cgataggctg	ctgtagttgg	gaagtgaagg	2220
tatcaggagt	tttttcttct	tcagaaagac	gttcgctcga	gaaaatcact	gaacggtttg	2280
cacttgtctg	aataactaaa	ggagatgcgg	tggtcggggt	gggagaagtc	acgctggctt	2340
gcaaagagga	gggaagaggc	gatttctttgc	tgctactttc	ttgtacaata	ggatcaaaaga	2400
aaagaatata	tccgtttgcga	gcttctaagg	aagaaagaat	cgcatctttc	tctaagtaga	2460
tggtggtttt	tactagacct	tggtcgagg	tggtttggga	gttattctgg	aacagaacag	2520
atccttcacc	tgcaaggata	gagagactcc	ctccggactg	gatggcgata	gctccacct	2580
ttttagcgct	gttggttaatg	aaggaaacag	gaccgttata	acgcaatacc	atgtgcttgg	2640
cataaatagc	cccgcctttt	tcttttagcag	aattgcttgc	aaatagcggt	tcttggttac	2700
aagcaagaga	gagagagcct	ttttctgagg	tggttagatc	gctgatacaa	caaatagcgc	2760
ctccattcgt	ggcggagttt	ccagtgaata	agagagggtt	tacattccct	gaacaaataa	2820
gattactaca	gcaaataagcg	ccgccatttta	aatccgcacg	attacggggc	aaagaaatag	2880
gcgacacatt	tctagataaaa	gagaagggtt	gagcctgaat	ggctccgcca	tttcttttag	2940
aagaattctc	ttcaaaagct	gtgaaaagat	agttgtgttg	tagagaaaag	gcatccgcag	3000
agatggctcc	tccagagcct	tcagcatgat	ttcgacagaa	actcatgcta	ccattattgc	3060
gcaaggataa	ctgaccattc	ttttgatgaa	taatagaagg	agagctttcc	ttggaagagc	3120
aatctgtgaa	ggaaaaggag	cgaaaatttt	gaatggaaag	atctcctttt	tctttaaaaa	3180
gagcctcttt	ggggtccgtg	atatggagtt	tagatacgaa	ccaacagtag	gaatttccaa	3240
gtacatacaa	aggatcatct	tgtgatgcat	gagcacaag	agcaggaaag	aggctctcca	3300
tggtacagtc	gtcaaaaaat	gtacagacga	catgattagg	attttttagt	aggagggcgg	3360
tttcacctaa	gggtaccctg	ccaaagagga	ccgctgtgga	tgacaaaata	gcagcacata	3420
gacaacagaa	gttcatatga	tcagggtcgca	tggtattccc	cctaagagtc	taaaagaaga	3480
ggttatgtca	aaaaaaaagta	ataatttaca	gactttttcc	tcaagggtct	tgtttcatgt	3540
ctttcaagat	gaagagttaa	ggaagatttt	tggtattgtg	aattagaaaa	atcagaatta	3600
tcaaattaga	tagtttttaa	aaaatttaaat	aatcttttaa	gagatcttat	tttattatcg	3660
atttttttta	ttttgatttt	taaaattaaa	ataagttttt	attttagttt	tgggttggtt	3720
tgttatgagc	atcaggggag	taggaggcaa	cggaatagt	cgaatccctt	ctcataatgg	3780
ggatggatcg	aatcgagaa	gtcaaaaata	gaagggtaat	aataaagttg	aagatcgagt	3840
ttgttctcta	tattcatctc	gtagtaacga	aaatagagaa	tctccttatg	cagtagtaga	3900
cgtcagctct	atgatcgaga	gcaccccaac	gagtggagag	acgacaagag	cttcgctgtg	3960
agtgttcagt	cgtttccaaa	gaggtttagt	acgagtagct	gacaaagtaa	gacgagctgt	4020
tcagtgtgcg	tggtgttcag	tctctacaag	aagatcgctc	gcaacaagag	ccgcagaatc	4080
cggatcaagt	agtcgtactg	ctcgtggtgc	aagttctggg	tatagggagt	attctccttc	4140
agcagctaga	gggtgtcgct	ttatgttcac	agatttctgg	agaactcggg	ttttacgcca	4200
gacctctcct	atggctggag	tttttgggaa	tcttgatgtg	aacgaggctc	gtttgatggc	4260
tgctgtacac	agtgatgcg	cggatcattt	agaagcgaac	aagttggctg	gccctgacgg	4320
ggtagcggcc	gcccgggaaa	ttgctaaaag	atgggagcaa	agagttagag	atctacaaga	4380
taaagggtgct	gcacgaaaat	tattaaatga	tccttttaggc	cgacgaacac	ctaattatca	4440
gagcaaaaat	ccagggtgagt	atactgtagg	gaattccatg	ttttacgatg	gtcctcagggt	4500
agcgaatctc	cagaacgtcg	acactggttt	ttggctggac	atgagcaatc	tctcagacgt	4560
tgtattatcc	agagagattc	aaacaggact	tcgagcacga	gctactttgg	aagaatccat	4620
gccgatgtta	gagaatttag	aagagcggtt	tagacgtttg	caagaaactt	gtgatgcggc	4680
tcgtactgag	atagaagaat	cgggatggac	tcgagagtc	gcatcaagaa	tggaaggcga	4740
tgaggcgcaa	ggaccttcta	gagcacaaca	agcttttcag	agctttgtaa	atgaatgtaa	4800
cagcatcgag	ttctcatttg	ggagctttgg	agagcatgtg	cgagttctct	gcgctagagt	4860
atcacgagga	ttagctgccg	caggagaggg	gattcgcctg	tgcttctctt	gttgtaaagg	4920
atcgacgcac	cgctacgctc	ctcgcgatga	cctatctcct	gaagggtgat	cgtagtagag	4980
gactttggct	agattcgcag	atgatattgg	aatagagcga	ggtgctgatg	gaacctacga	5040
tattcctttg	gtagatgatt	ggagaagagg	ggttcctagt	attgaaggag	aaggatctga	5100
ctcgatctat	gaaatcatga	tgcttatcta	tgaagttatg	gatattggatc	tagaaacacg	5160
aagatctttt	gcggtacagc	aagggcacta	tcaggaccca	agagcttcag	attatgacct	5220
cccacgtgct	agegactatg	atttgcctag	aagcccatat	cctactccac	ctttgcctcc	5280
tagatatcag	ctacagaata	tggtatgtaga	agcagggttc	cgtgaggcag	tttatgcttc	5340
ttttgtagca	ggaatgtaca	attatgtagt	gacacagccg	caagagcgta	ttcccaatag	5400
tcagcaggtg	gaagggtatc	tgctgtatat	gcttaccac	gggtcacaga	catttagaga	5460
cctgatgagg	cgttggaata	gagaagtcga	tagggaataa	actggtatct	accataggtt	5520

tgtagcaaaa aactaagccc accaagaaga aattctcttt ggtgggcttc tttttttatt  
caaaaaagaa agccctcttc a

5580  
5601

//

PAGE BLANK (USPTO)



```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 08:37:27 ON 31 OCT 2001
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2001 AMERICAN CHEMICAL SOCIETY (ACS)
```

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications.

```
FILE COVERS 1947 - 31 Oct 2001 VOL 135 ISS 19
FILE LAST UPDATED: 30 Oct 2001 (20011030/ED)
```

This file contains CAS Registry Numbers for easy and accurate substance identification.

This file supports REGISTRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

HCAplus now provides online access to patents and literature covered in CA from 1947 to the present. On April 22, 2001, bibliographic information and abstracts were added for over 2.2 million references published in CA from 1947 to 1966.

=>

=>

```
=> d stat que 15
L1      55 SEA FILE=HCAPLUS ABB=ON PLU=ON ("PROBST P"/AU OR "PROBST
      P"/IN OR "PROBST P A"/AU OR "PROBST P J"/AU) OR ("PROBST
      PETER"/AU OR "PROBST PETER"/IN OR "PROBST PETER G"/AU OR
      "PROBST PETER J"/AU OR "PROBST PETER S"/AU)
L2      356 SEA FILE=HCAPLUS ABB=ON PLU=ON ("BHATIA A"/AU OR "BHATIA A
      B"/AU OR "BHATIA A K"/AU OR "BHATIA A L"/AU OR "BHATIA A S"/AU
      OR "BHATIA AJAY"/AU OR "BHATIA AJAY"/IN)
L3      70 SEA FILE=HCAPLUS ABB=ON PLU=ON ("SKEIKY Y A W"/AU OR "SKEIKY
      YASIR"/AU OR "SKEIKY YASIR"/IN OR "SKEIKY YASIR A"/AU OR
      "SKEIKY YASIR A"/IN OR "SKEIKY YASIR A W"/AU OR "SKEIKY YASIR
      A W"/IN)
L4      22 SEA FILE=HCAPLUS ABB=ON PLU=ON ("FLING STEVEN"/AU OR "FLING
      STEVEN P"/AU OR "FLING STEVEN P"/IN OR "FLING STEVEN PAUL"/AU)
L5      2 SEA FILE=HCAPLUS ABB=ON PLU=ON L1 AND L2 AND L3 AND L4
```

=>

=>

```
=> d ibib abs hitrn 15 1-2
```

```
L5 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2001 ACS
ACCESSION NUMBER: 2001:417155 HCAPLUS
DOCUMENT NUMBER: 135:45174
TITLE: Antigenic compounds and methods for treatment and
      diagnosis of Chlamydial infection
```

**THIS PAGE BLANK (USPTO)**

INVENTOR(S): **Probst, Peter**; Bhatia, Ajay; Skeiky, Yasir  
 A. W.; Fling, Steven P.; Scholler, John  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 293 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001040474	A2	20010607	WO 2000-US32919	20001204
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
PRIORITY APPLN. INFO.:			US 1999-454684	A 19991203
			US 2000-556877	A 20000419
			US 2000-598419	A 20000620
AB Compds. and methods for the diagnosis and treatment of Chlamydial infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a Chlamydia antigen and DNA sequences encoding such polypeptides from Chlamydia trachomatis and C. pneumoniae isolated using retroviral expression vector systems and subsequent immunol. anal. and epitope mapping. Pharmaceutical compns. and vaccines comprising such polypeptides or DNA sequences are also provided, together with antibodies directed against such polypeptides. Diagnostic kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of Chlamydial infection in patients and in biol. samples. In particular, fusion proteins are constructed from the Chlamydial proteins PmpA, PmpF, PmpH, PmpB, and PmpC fused with amino acid residues 192-323 of the Ra2 MTB32A serine proteinase from Mycobacterium tuberculosis.				

L5 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:402007 HCAPLUS

DOCUMENT NUMBER: 133:53686

TITLE: Chlamydial antigens and genomic DNA sequences for  
 treatment and diagnosis of chlamydial infection

INVENTOR(S): **Probst, Peter**; Bhatia, Ajay; Skeiky, Yasir  
 A. W.; Fling, Steven P.; Jen, Shyian; Stromberg, Erica  
 Jean

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: PCT Int. Appl., 256 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000034483	A2	20000615	WO 1999-US29012	19991208

**THIS PAGE BLANK (USPTO)**

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

US 6166177 A 20001226 US 1998-208277 19981208

EP 1144642 A2 20011017 EP 1999-963037 19991208

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO

NO 2001002812 A 20010802 NO 2001-2812 20010607

PRIORITY APPLN. INFO.:

US 1998-208277 A 19981208

US 1999-288594 A 19990408

US 1999-410568 A 19991001

US 1999-426571 A 19991022

WO 1999-US29012 W 19991208

AB Compds. and methods for the diagnosis and treatment of Chlamydial infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a Chlamydia antigen and DNA sequences encoding such polypeptides. Chlamydia antigens were isolated by expression cloning of a genomic DNA library of C. trachomatis LGV II, and shown to induce T cell proliferation and interferon-.beta. prodn. Immune responses of human PBMC and T cell lines are generated against the Chlamydia antigens. Pharmaceutical compns. and vaccines comprising such polypeptides or DNA sequences are also provided, together with antibodies directed against such polypeptides. Diagnostic kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of Chlamydial infection in patients and in biol. samples.

=>

=>

=> d stat que 16

L1 55 SEA FILE=HCAPLUS ABB=ON PLU=ON ("PROBST P"/AU OR "PROBST P"/IN OR "PROBST P A"/AU OR "PROBST P J"/AU) OR ("PROBST PETER"/AU OR "PROBST PETER"/IN OR "PROBST PETER G"/AU OR "PROBST PETER J"/AU OR "PROBST PETER S"/AU)

L2 356 SEA FILE=HCAPLUS ABB=ON PLU=ON ("BHATIA A"/AU OR "BHATIA A B"/AU OR "BHATIA A K"/AU OR "BHATIA A L"/AU OR "BHATIA A S"/AU OR "BHATIA AJAY"/AU OR "BHATIA AJAY"/IN)

L3 70 SEA FILE=HCAPLUS ABB=ON PLU=ON ("SKEIKY Y A W"/AU OR "SKEIKY YASIR"/AU OR "SKEIKY YASIR"/IN OR "SKEIKY YASIR A"/AU OR "SKEIKY YASIR A"/IN OR "SKEIKY YASIR A W"/AU OR "SKEIKY YASIR A W"/IN)

L4 22 SEA FILE=HCAPLUS ABB=ON PLU=ON ("FLING STEVEN"/AU OR "FLING STEVEN P"/AU OR "FLING STEVEN P"/IN OR "FLING STEVEN PAUL"/AU)

L5 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L1 AND L2 AND L3 AND L4

L6 3 SEA FILE=HCAPLUS ABB=ON PLU=ON (L1 AND (L2 OR L3 OR L4)) NOT L5

=>

=>

**THIS PAGE BLANK (USPTO)**

=&gt; d ibib abs hitrn 16 1-3

L6 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2001:780955 HCAPLUS  
 TITLE: Leishmania antigens for use in the therapy and  
 diagnosis of leishmaniasis  
 INVENTOR(S): Reed, Steven G.; Campos-Neto, Antonio; Webb, John R.;  
 Dillon, Davin C.; **Skeiky, Yasir A. W.**;  
**Bhatia, Ajay**; Coler, Rhea M.; **Probst,**  
**Peter**  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 193 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001079276	A2	20011025	WO 2001-US11254	20010405
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
PRIORITY APPLN. INFO.:			US 2000-551974	A 20000414
			US 2000-565501	A 20000505
			US 2000-639206	A 20000814
AB Compositions and methods for preventing, treating and detecting leishmaniasis and stimulating immune responses in patients are disclosed. The compounds provided include polypeptides that contain at least an immunogenic portion of one or more Leishmania antigens, or a variant thereof. Vaccines and pharmaceutical compositions comprising such polypeptides, or polynucleotides encoding such polypeptides, are also provided and may be used, for example, for the prevention and therapy of leishmaniasis, as well as for the detection of Leishmania infection.				

L6 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2001:120546 HCAPLUS  
 DOCUMENT NUMBER: 134:325125  
 TITLE: CD8+ T cells recognize an inclusion  
 membrane-associated protein from the vacuolar pathogen  
 Chlamydia trachomatis  
 AUTHOR(S): **Fling, Steven P.**; Sutherland, R. Alec;  
 Steele, Lisa N.; Hess, Bruce; D'Orazio, Sarah E. F.;  
 Maisonneuve, Jean-Francois; Lampe, Mary F.;  
**Probst, Peter**; Starnbach, Michael N.  
 CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA  
 SOURCE: Proc. Natl. Acad. Sci. U. S. A. (2001), 98(3),  
 1160-1165  
 CODEN: PNASA6; ISSN: 0027-8424  
 PUBLISHER: National Academy of Sciences  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English

**THIS PAGE BLANK (USPTO)**



AB During infection with *Chlamydia trachomatis*, CD8+ T cells are primed, even though the bacteria remain confined to a host cell vacuole throughout their developmental cycle. Because CD8+ T cells recognize antigens processed from cytosolic proteins, the *Chlamydia* antigens recognized by these CD8+ T cells very likely have access to the host cell cytoplasm during infection. The identity of these *C. trachomatis* proteins has remained elusive, even though their localization suggests they may play important roles in the biol. of the organism. Here we use a retroviral expression system to identify Cap1, a 31-kDa protein from *C. trachomatis* recognized by protective CD8+ T cells. Cap1 contains no strong homol. to any known protein. Immunofluorescence microscopy by using Cap1-specific antibody demonstrates that this protein is localized to the vacuolar membrane. Cap1 is virtually identical among the human *C. trachomatis* serovars, suggesting that a vaccine incorporating Cap1 might enable the vaccine to protect against all *C. trachomatis* serovars. The identification of proteins such as Cap1 that assoc. with the inclusion membrane will be required to fully understand the interaction of *C. trachomatis* with its host cell.

REFERENCE COUNT: 34

REFERENCE(S): (1) Bannantine, J; Mol Microbiol 1998, V28, P1017 HCAPLUS  
 (2) Blattner, J; J Cell Biol 1992, V119, P1129 HCAPLUS  
 (4) Chakrabarti, S; Mol Cell Biol 1985, V5, P3403 HCAPLUS  
 (5) Denamur, E; J Gen Microbiol 1991, V137, P2525 HCAPLUS  
 (6) Earl, P; J Virol 1991, V65, P31 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1997:692504 HCAPLUS

DOCUMENT NUMBER: 127:306395

TITLE: A Leishmania protein that modulates interleukin (IL)-12, IL-10, and tumor necrosis factor-.alpha. production and expression of B7-1 in human monocyte-derived antigen-presenting cells

AUTHOR(S): Probst, Peter; Skeiky, Yasir A. W.  
 ; Steeves, Michelle; Gervassi, Ana; Grabstein, Kenneth H.; Reed, Steven G.

CORPORATE SOURCE: Infectious Disease Research Inst., Seattle, WA, USA

SOURCE: Eur. J. Immunol. (1997), 27(10), 2634-2642

CODEN: EJIMAF; ISSN: 0014-2980

PUBLISHER: Wiley-VCH

DOCUMENT TYPE: Journal

LANGUAGE: English

AB LeIF, a gene homolog of the eukaryotic initiation factor 4A was 1st described as a leishmanial antigen that induced a Th1-type T cell response in peripheral blood mononuclear cells (PBMC) from leishmaniasis patients. The interferon (IFN)-.gamma. prodn. by PBMC was interleukin (IL)-12 dependent. The effects of LeIF were characterized on cytokine prodn. and expression of surface mols. by normal human monocytes as well as by monocyte-derived macrophages and dendritic cells (MoDC); LeIF was a strong inducer of IL-12 and, to a lesser extent, of IL-10, and tumor necrosis factor (TNF)-.alpha. in macrophages and MoDC. IL-12 prodn. did not require CD40 triggering, confirming that the ability of LeIF to induce IL-12 was, not mediated through an effect on T cells. However, addn. of sol. CD40 ligand (L) synergistically augmented IL-12 prodn. in macrophages and MoDC. The cytokine-inducing activity of LeIF is located in the N-terminal portion of the mol. and was both proteinase K sensitive and

**THIS PAGE BLANK (USPTO)**

polymyxin B resistant. LeIF, lipopolysaccharide, and fixed *Staphylococcus aureus* all induced comparable amts. of IL-12, validating the potent cytokine-inducing effects of LeIF. Of these stimuli, LeIF had the highest IL-12/IL-10 and IL-12/TNF- $\alpha$  ratio demonstrating the preference of LeIF for IL-12 induction. Studies investigating the expression of surface mols. showed that LeIF up-regulated B7-1 and CD54 (ICAM-1) on macrophages and MoDC. This is the 1st report describing IL-12 prodn., up-regulation of co-stimulatory and intercellular adhesion mols. by monocytic antigen-presenting cells in response to a protein from a pathogenic microorganism. These immunomodulatory characteristics of LeIF might be excellent properties for a Th11-type adjuvant.

=&gt;

=&gt;

=&gt; d stat que 17

```

L1      55 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("PROBST P"/AU OR "PROBST
        P"/IN OR "PROBST P A"/AU OR "PROBST P J"/AU) OR ("PROBST
        PETER"/AU OR "PROBST PETER"/IN OR "PROBST PETER G"/AU OR
        "PROBST PETER J"/AU OR "PROBST PETER S"/AU)
L2      356 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("BHATIA A"/AU OR "BHATIA A
        B"/AU OR "BHATIA A K"/AU OR "BHATIA A L"/AU OR "BHATIA A S"/AU
        OR "BHATIA AJAY"/AU OR "BHATIA AJAY"/IN)
L3      70 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("SKEIKY Y A W"/AU OR "SKEIKY
        YASIR"/AU OR "SKEIKY YASIR"/IN OR "SKEIKY YASIR A"/AU OR
        "SKEIKY YASIR A"/IN OR "SKEIKY YASIR A W"/AU OR "SKEIKY YASIR
        A W"/IN)
L4      22 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("FLING STEVEN"/AU OR "FLING
        STEVEN P"/AU OR "FLING STEVEN P"/IN OR "FLING STEVEN PAUL"/AU)
L5      2 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L1 AND L2 AND L3 AND L4
L6      3 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L1 AND (L2 OR L3 OR L4)) NOT
        L5
L7      4 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L2 AND (L3 OR L4)) NOT (L5
        OR L6)

```

=&gt;

=&gt;

=&gt; d ibib abs hitrn 17 1-4

L7 ANSWER 1 OF 4 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1999:534897 HCAPLUS

DOCUMENT NUMBER: 131:334383

TITLE: Cloning, characterization and serological evaluation  
of K9 and K26: two related hydrophilic antigens of  
*Leishmania chagasi*

AUTHOR(S): **Bhatia, A.**; Daifalla, N. S.; Jen, S.;  
Badaro, R.; Reed, S. G.; **Skeiky, Y. A. W.**

CORPORATE SOURCE: Corixa Corporation, Seattle, WA, USA

SOURCE: Mol. Biochem. Parasitol. (1999), 102(2), 249-261

CODEN: MBIPDP; ISSN: 0166-6851

PUBLISHER: Elsevier Science Ireland Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB We report here the mol. cloning and characterization of two related  
hydrophilic antigens of *Leishmania chagasi*. These two antigens have  
predicted mol. wts. of .apprx.9 and 26 kDa and detect antibodies in sera  
of patients with kala-azar (k). Thus, to maintain consistency with

**THIS PAGE BLANK (USPTO)**

nomenclature of the previously described 39kDa diagnostic antigen of *L. chagasi* (k39 [1]), these antigens are being referred to as k9 and k26. A significant difference between k9 and k26 is the presence of 11 copies of a 14 amino acid repeat in the open reading frame of k26. The region flanking the repeats of k26 shares a 69% identity with the open reading frame of k9. The recombinant proteins encoded by both antigens are very hydrophilic and show aberrant migration on SDS PAGE. Results of Southern blot anal. reveal that k9 and k26 are conserved to varying degrees among various *Leishmania* species. Interestingly, the repeat region of k26 is specific to *L. chagasi* and *L. donovani* while the flanking region is conserved among several other species. Transcript levels of k26 are significantly upregulated in the amastigote stage of the parasite. Our results show that recombinant K26 is specific in detecting antibodies in infection sera from visceral leishmaniasis (VL) patients. Thus rK26 may complement rK39 in a more accurate diagnosis of VL in the old and the new world.

REFERENCE COUNT: 29  
 REFERENCE(S): (1) Aly, R; Nucl Acids Res 1994, V22, P2922 HCAPLUS  
 (3) Badaro, R; J Infect Dis 1996, V173, P758 HCAPLUS  
 (4) Burns, J; Proc Natl Acad Sci USA 1993, V90, P775 HCAPLUS  
 (5) Chomczynski, P; Anal Biochem 1987, V162, P156 HCAPLUS  
 (6) Feinberg, A; Anal Biochem 1983, V132, P6 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 2 OF 4 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1997:415380 HCAPLUS

DOCUMENT NUMBER: 127:91226

TITLE: Molecular cloning and characterization of two iron superoxide dismutase cDNAs from *Trypanosoma cruzi*  
 AUTHOR(S): Ismail, Said O.; Paramchuk, Wendy; **Skeiky, Yasir A. W.**; Reed, Steven G.; **Bhatia, Ajay**;  
 Gedamu, Lashitew

CORPORATE SOURCE: Dep. Biological Sciences, Univ. Calgary, Calgary, AB, T2N 1N4, Can.

SOURCE: Mol. Biochem. Parasitol. (1997), 86(2), 187-197  
 CODEN: MBIPDP; ISSN: 0166-6851

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Two cDNAs (FeSODA and FeSODB cDNAs) corresponding to superoxide dismutase (1.15.1.1., SOD) were isolated from a *Trypanosoma cruzi* cDNA library. Comparison of the deduced amino acid sequences with previously reported SOD protein sequences revealed that the *T. cruzi* open reading frames had considerable homol. with FeSODs. The coding region of the *T. cruzi* FeSODB cDNA has been expressed in fusion with glutathione-S-transferase using an *Escherichia coli* mutant QC779, lacking both MnSOD and FeSOD genes (sodA sodB). Staining of native polyacrylamide gels for SOD activity of *T. cruzi* crude lysate and the recombinant SOD suggests that this protein is an FeSOD. The recombinant enzyme also protected the *E. coli* mutant QC779 from paraquat toxicity. Northern blot anal. showed that FeSODB is differentially expressed, showing a higher level at the epimastigote stage of *T. cruzi* development; whereas, FeSODA is constitutively expressed at a lower level in all developmental stages. Furthermore, Southern hybridization shows that both FeSODA and FeSODB genes appear to be present in the *T. cruzi* genome as multiple repeating units (multi-copy gene family).

**THIS PAGE BLANK (USPTO)**

L7 ANSWER 3 OF 4 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1995:781475 HCAPLUS  
 DOCUMENT NUMBER: 123:277157  
 TITLE: Molecular cloning, characterization, and expression in Escherichia coli of iron superoxide dismutase cDNA from Leishmania donovani chagasi. [Retraction to document cited in CA120:184186]  
 AUTHOR(S): Ismail, Said O.; **Skeiky, Yasir A. W.**; **Bhatia, Ajay**; Omara-Opyene, Levi A.; Gedamu, Lashitew  
 CORPORATE SOURCE: Dep. Biol. Sci., Univ. Calgary, Calgary, AB, T2N 1N4, Can.  
 SOURCE: Infect. Immun. (1995), 63(9), 3749  
 CODEN: INFIBR; ISSN: 0019-9567  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English  
 AB The authors have requested that the article be retracted, as a Trypanosoma cruzi dDNA library rather than a Leishmania donovani chagasi dDNA library was used in screening.

L7 ANSWER 4 OF 4 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1994:184186 HCAPLUS  
 DOCUMENT NUMBER: 120:184186  
 TITLE: Molecular cloning, characterization, and expression in Escherichia coli of iron superoxide dismutase cDNA from Leishmania donovani chagasi  
 AUTHOR(S): Ismail, Said O.; **Skeiky, Yasir A. W.**; **Bhatia, Ajay**; Omara-Opyene, Levi A.; Gedamu, Lashitew  
 CORPORATE SOURCE: Dep. Biol. Sci., Univ. Calgary, Calgary, AB, T2N 1N4, Can.  
 SOURCE: Infect. Immun. (1994), 62(2), 657-64  
 CODEN: INFIBR; ISSN: 0019-9567  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English  
 AB A cDNA corresponding to superoxide dismutase (SOD; EC 1.15.1.1.) was isolated from a Leishmania donovani chagasi (L. d. chagasi) promastigote cDNA library, using PCR with a set of primers derived from conserved amino acids of manganese SODs (MnSODs) and iron SODs (FeSODs). Comparison of the deduced amino acid sequences with previously reported SOD amino acid sequences revealed that the L. d. chagasi 585-bp open reading frame had considerable homol. with FeSODs and MnSODs. The highest homol. was shared with prokaryotic FeSODs. The coding region of L. d. chagasi SOD cDNA has been expressed in fusion with glutathione-S-transferase, using an Escherichia coli mutant, QC779, lacking both MnSOD and FeSOD genes (sodA and sodB). Staining of native polyacrylamide gels for SOD activity of Leishmania crude lysate and the recombinant SOD revealed that both had SOD activity that was inactivated by 5 mM hydrogen peroxide but not by 2 mM potassium cyanide, which is indicative of FeSOD. The recombinant enzyme also protected E. coli mutant QC779 from paraquat toxicity. This indicated that the glutathione-S-transferase peptide does not interfere with the in vivo and in vitro activities of the recombinant SOD. Cross-species hybridization showed that FeSOD is highly conserved in the Leishmania genus. Interestingly, the hybridization pattern of the FeSOD gene(s) coincided with other classification schemes that divide Leishmania species into complexes. The cloning of FeSOD cDNA may contribute to the understanding of the role of SODs in Leishmania pathogenesis.

**THIS PAGE BLANK (USPTO)**



=&gt;

=&gt;

=&gt; d stat que 18

```

L1      55 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("PROBST P"/AU OR "PROBST
P"/IN OR "PROBST P A"/AU OR "PROBST P J"/AU) OR ("PROBST
PETER"/AU OR "PROBST PETER"/IN OR "PROBST PETER G"/AU OR
"PROBST PETER J"/AU OR "PROBST PETER S"/AU)
L2      356 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("BHATIA A"/AU OR "BHATIA A
B"/AU OR "BHATIA A K"/AU OR "BHATIA A L"/AU OR "BHATIA A S"/AU
OR "BHATIA AJAY"/AU OR "BHATIA AJAY"/IN)
L3      70 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("SKEIKY Y A W"/AU OR "SKEIKY
YASIR"/AU OR "SKEIKY YASIR"/IN OR "SKEIKY YASIR A"/AU OR
"SKEIKY YASIR A"/IN OR "SKEIKY YASIR A W"/AU OR "SKEIKY YASIR
A W"/IN)
L4      22 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("FLING STEVEN"/AU OR "FLING
STEVEN P"/AU OR "FLING STEVEN P"/IN OR "FLING STEVEN PAUL"/AU)
L5      2 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L1 AND L2 AND L3 AND L4
L6      3 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L1 AND (L2 OR L3 OR L4)) NOT
L5
L8      0 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L3 AND L4) NOT (L5 OR L6)

```

=&gt; d stat que 112

```

L1      55 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("PROBST P"/AU OR "PROBST
P"/IN OR "PROBST P A"/AU OR "PROBST P J"/AU) OR ("PROBST
PETER"/AU OR "PROBST PETER"/IN OR "PROBST PETER G"/AU OR
"PROBST PETER J"/AU OR "PROBST PETER S"/AU)
L2      356 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("BHATIA A"/AU OR "BHATIA A
B"/AU OR "BHATIA A K"/AU OR "BHATIA A L"/AU OR "BHATIA A S"/AU
OR "BHATIA AJAY"/AU OR "BHATIA AJAY"/IN)
L3      70 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("SKEIKY Y A W"/AU OR "SKEIKY
YASIR"/AU OR "SKEIKY YASIR"/IN OR "SKEIKY YASIR A"/AU OR
"SKEIKY YASIR A"/IN OR "SKEIKY YASIR A W"/AU OR "SKEIKY YASIR
A W"/IN)
L4      22 SEA FILE=HCAPLUS ABB=ON  PLU=ON  ("FLING STEVEN"/AU OR "FLING
STEVEN P"/AU OR "FLING STEVEN P"/IN OR "FLING STEVEN PAUL"/AU)
L5      2 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L1 AND L2 AND L3 AND L4
L6      3 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L1 AND (L2 OR L3 OR L4)) NOT
L5
L7      4 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L2 AND (L3 OR L4)) NOT (L5
OR L6)
L11     30 SEA FILE=HCAPLUS ABB=ON  PLU=ON  (L1 OR L2 OR L3 OR L4 OR L5)
AND ((INFECT? OR ?CHLAMYD? OR BACTER? OR VIRUS?) (L) (DIAGNOSIS
OR DX OR DIAGNOS?))
L12     27 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L11 NOT (L5 OR L6 OR L7)

```

=&gt;

=&gt;

=&gt; d ibib abs hitrn 112 1-27

L12 ANSWER 1 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:687327 HCAPLUS

DOCUMENT NUMBER: 135:271872

TITLE: Compounds and methods for immunotherapy and diagnosis  
of tuberculosis

INVENTOR(S): Reed, Steven G.; Skeiky, Yasir A. W.;

**THIS PAGE BLANK (USPTO)**

PATENT ASSIGNEE(S): Dillon, Davin C.; Campos-Neto, Antonio; Houghton, Raymond; Vedvick, Thomas S.; Twardzik, Daniel R.  
 SOURCE: Corixa Corporation, USA  
 U.S., 100 pp., Cont.-in-part of U.S. Ser. No. 730,510.  
 CODEN: USXXAM  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 3  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6290969	B1	20010918	US 1997-818112	19970313
ZA 9607394	A	19970505	ZA 1996-7394	19960830
WO 9816646	A2	19980423	WO 1997-US18293	19971007
WO 9816646	A3	19981008		
W: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
AU 9748144	A1	19980511	AU 1997-48144	19971007
EP 932681	A2	19990804	EP 1997-910873	19971007
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
BR 9712518	A	20001024	BR 1997-12518	19971007
JP 2001501832	T2	20010213	JP 1998-518456	19971007
NO 9901694	A	19990610	NO 1999-1694	19990409
PRIORITY APPLN. INFO.:				
US 1995-523436 B2 19950901				
US 1995-533634 B2 19950922				
US 1996-620874 B2 19960322				
US 1996-659683 A2 19960605				
US 1996-680574 A2 19960712				
US 1996-730510 A2 19961011				
US 1997-818112 A 19970313				
WO 1997-US18293 W 19971007				
AB Comps. and methods for inducing protective immunity against tuberculosis are disclosed. The comps. provided include polypeptides that contain at least one immunogenic portion of one or more Mycobacterium tuberculosis proteins and DNA mols. encoding such polypeptides. Such comps. may be formulated into vaccines and/or pharmaceutical comps. for immunization against M. tuberculosis <b>infection</b> , or may be used for the <b>diagnosis</b> of tuberculosis.				
REFERENCE COUNT: 61				
REFERENCE(S):				
(2) Andersen; Infection and Immunity 1993, V61(3), P844 HCAPLUS				
(3) Andersen; Scand J Immunol 1992, V36, P823 HCAPLUS				
(4) Andersen, P; Infection and Immunity 1994, V62(6), P2536 HCAPLUS				
(5) Anon; WO 8805823 1988 HCAPLUS				
(6) Anon; EP 419355 A1 1991 HCAPLUS				
ALL CITATIONS AVAILABLE IN THE RE FORMAT				
L12 ANSWER 2 OF 27 HCAPLUS COPYRIGHT 2001 ACS				
ACCESSION NUMBER: 2001:636186 HCAPLUS				
DOCUMENT NUMBER: 135:209890				

**THIS PAGE BLANK (USPTO)**

TITLE: Compounds and methods for diagnosis and immunotherapy of tuberculosis  
 INVENTOR(S): Campos-Neto, Antonio; **Skeiky, Yasir**; Ovendale, Pamela; Jen, Shyian; Lodes, Michael  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 161 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001062893	A2	20010830	WO 2001-US5992	20010226
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
PRIORITY APPLN. INFO.:			US 2000-185037	P 20000225
			US 2000-223828	P 20000808

AB Compsds. and methods for **diagnosing** tuberculosis or for inducing protective immunity against tuberculosis are disclosed. The compds. provided include polypeptides that contain at least one immunogenic portion of one or more Mycobacterium proteins and DNA mols. encoding such polypeptides. **Diagnostic** kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of Mycobacterium **infection** in patients and biol. samples. Antibodies directed against such polypeptides are also provided. In addn., such compds. may be formulated into vaccines and/or pharmaceutical compns. for immunization against Mycobacterium **infection**.

L12 ANSWER 3 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2001:582322 HCAPLUS  
 DOCUMENT NUMBER: 135:179697  
 TITLE: Diagnosis of tuberculosis  
 INVENTOR(S): Alderson, Mark R.; Dillon, Davin C.; **Skeiky, Yasir A. w.**; Campos-Neto, Antonio  
 PATENT ASSIGNEE(S): USA  
 SOURCE: U.S. Pat. Appl. Publ., 74 pp., Cont.-in-part of U.S. Ser. No. 858,998, abandoned.  
 CODEN: USXXCO  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 2  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2001012888	A1	20010809	US 1998-73009	19980505
WO 9853076	A2	19981126	WO 1998-US10514	19980520
WO 9853076	A3	19990401		
W: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP,				

**THIS PAGE BLANK (USPTO)**

KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO,  
 NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA,  
 UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM  
 RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES,  
 FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI,  
 CM, GA, GN, ML, MR, NE, SN, TD, TG

AU 9875916	A1	19981211	AU 1998-75916	19980520
BR 9809443	A	20000613	BR 1998-9443	19980520
NO 9905690	A	20000118	NO 1999-5690	19991119
PRIORITY APPLN. INFO.:			US 1997-858998	B2 19970520
			US 1998-73009	A 19980505
			WO 1998-US10514	W 19980520

AB Compds. and methods for **diagnosing** tuberculosis are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of one or more Mycobacterium tuberculosis proteins, and DNA sequences encoding such polypeptides. **Diagnostic** kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of M. tuberculosis **infection** in patients and biol. samples. Antibodies directed against such polypeptides are also provided.

L12 ANSWER 4 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:550008 HCAPLUS  
 TITLE: Serological expression cloning and immunological evaluation of MTB48, a novel mycobacterium tuberculosis antigen  
 AUTHOR(S): Lodes, Michael J.; Dillon, Davin C.; Mohamath, Raodoh; Day, Craig H.; Benson, Darin R.; Reynolds, Lisa D.; McNeill, Patricia; Sampaio, Diana Pedral; **Skeiky, Yasir A. W.**; Badaro, Roberto; Persing, David H.; Reed, Steven G.; Houghton, Raymond L.  
 CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA  
 SOURCE: J. Clin. Microbiol. (2001), 39(7), 2485-2493  
 CODEN: JCMIDW; ISSN: 0095-1137  
 PUBLISHER: American Society for Microbiology  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English

AB Improved **diagnostics** are needed for the detection of Mycobacterium tuberculosis, esp. for patients with smear-neg. disease. To address this problem, we have screened M. tuberculosis (H37Rv and Erdman strains) genomic expression libraries with pooled sera from patients with extrapulmonary disease and with sera from patients with elevated reactivity with M. tuberculosis lysate. Both serum pools were reactive with clones expressing a recombinant protein referred to here as MTB48. The genomic sequence of the resulting clones was identical to that of the M. tuberculosis H37Rv isolate and showed 99% identity to the Mycobacterium bovis and M. bovis BCG isolate sequences. The genomic location of this sequence is 826 bp upstream of a region contg. the esat-6 gene that is deleted in the M. bovis BCG isolate. The mtb48 1,380-bp open reading frame encodes a predicted 47.6-kDa polypeptide with no known function. Southern and Western blot analyses indicate that this sequence is present in a single copy and is conserved in the M. tuberculosis and M. bovis isolates tested but not in other mycobacterial species tested, including Mycobacterium leprae and Mycobacterium avium. In addn., the native protein was detected in the cytoplasm, as was a processed form that was also shed into the medium during culture. Serol. anal. of recombinant MTB48 and the M. tuberculosis 38-kDa antigen with a panel of patient and control sera indicates that the inclusion of recombinant MTB48 in a prototype serodiagnostic test increases assay sensitivity for M.

**THIS PAGE BLANK (USPTO)**



tuberculosis **infection** when it is combined with other known immunodominant antigens, such as the 39-kDa antigen.

REFERENCE COUNT: 35  
 REFERENCE(S): (1) Andersen, A; Infect Immun 1989, V57, P2481 HCAPLUS  
 (2) Ashbridge, K; J Immunol 1990, V144, P3137 HCAPLUS  
 (9) Dillon, D; J Clin Microbiol 2000, V38, P3285 HCAPLUS  
 (18) Harboe, M; J Infect Dis 1992, V166, P874 HCAPLUS  
 (20) Hendrickson, R; J Clin Microbiol 2000, V38, P2354 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 5 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2001:331212 HCAPLUS  
 DOCUMENT NUMBER: 134:352261  
 TITLE: Compounds and methods for the detection and prevention of T. cruzi infection  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Lodes, Michael J.; Houghton, Raymond L.; Smith, John M.; McNeill, Patricia D.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: U.S., 56 pp., Cont.-in-part of U.S. 6,054,135.  
 CODEN: USXXAM  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 4  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6228372	B1	20010508	US 1997-993674	19971218
US 6054135	A	20000425	US 1997-834306	19970415
WO 9931246	A1	19990624	WO 1998-US25871	19981204
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 9918057	A1	19990705	AU 1999-18057	19981204
EP 1038000	A1	20000927	EP 1998-962923	19981204
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				

PRIORITY APPLN. INFO.: US 1997-834306 A2 19970415  
 WO 1996-US18624 A2 19961114  
 US 1997-993674 A 19971218  
 WO 1998-US25871 W 19981204

AB Compds. and methods are provided for **diagnosing** Trypanosoma cruzi **infection**. The disclosed compds. are polypeptides, or antibodies thereto, that contain one or more epitopes of T. cruzi antigens. The compds. are useful in a variety of immunoassays for detecting T. cruzi **infection**. The polypeptide compds. are further useful in vaccines and pharmaceutical compns. for inducing protective immunity against Chagas' disease in individuals exposed to T. cruzi.

REFERENCE COUNT: 23  
 REFERENCE(S): (2) Anon; WO 9209895 1992 HCAPLUS

**THIS PAGE BLANK (USPTO)**

(3) Anon; WO 9316199 1993 HCAPLUS  
 (4) Anon; WO 9401776 1994 HCAPLUS  
 (5) Anon; WO 9629605 1996 HCAPLUS  
 (6) Anon; WO 9718475 1997 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 6 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2001:265269 HCAPLUS  
 DOCUMENT NUMBER: 134:309685  
 TITLE: Fusion proteins of Mycobacterium tuberculosis  
 INVENTOR(S): **Skeiky, Yasir**; Reed, Steven; Houghton,  
 Raymond L.; Mcneill, Patricia D.; Dillon, Davin C.;  
 Lodes, Michael L.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 168 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001024820	A1	20010412	WO 2000-US28095	20001010
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
PRIORITY APPLN. INFO.:			US 1999-158338	P 19991007
			US 1999-158425	P 19991007

AB The present invention relates to fusion proteins contg. at least two Mycobacterium species antigens. In particular, it relates to nucleic acids encoding fusion proteins that include two or more individual M. tuberculosis antigens, which increase serol. sensitivity of sera from individuals **infected** with tuberculosis, and methods for their use in the **diagnosis**, treatment, and prevention of tuberculosis **infection**.

REFERENCE COUNT: 9  
 REFERENCE(S): (1) Alderson; J Exp Med 2000, V191(3), P551 HCAPLUS  
 (2) Brandt; Infect Immun 2000, V68(2), P791 HCAPLUS  
 (3) Coler; J Immunol 1998, V161(5), P2356 HCAPLUS  
 (4) Corixa Corporation; WO 9951748 A2 1999 HCAPLUS  
 (5) Hendrickson; J Clin Microbiol 2000, V38(6), P2354 HCAPLUS

ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 7 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2000:698208 HCAPLUS  
 DOCUMENT NUMBER: 134:235797  
 TITLE: Molecular and immunological characterization of Mycobacterium tuberculosis CFP-10, an immunodiagnostic antigen missing in Mycobacterium bovis BCG  
 AUTHOR(S): Dillon, Davin C.; Alderson, Mark R.; Day, Craig H.; Bement, Teresa; Campos-Neto, Antonio; **Skeiky**,

**THIS PAGE BLANK (USPTO)**

CORPORATE SOURCE: Yasir A. W.; Vedvick, Thomas; Badaro, Roberto;  
SOURCE: Reed, Steven G.; Houghton, Raymond  
Corixa Corporation, Seattle, WA, 98104, USA  
J. Clin. Microbiol. (2000), 38(9), 3285-3290  
CODEN: JCMIDW; ISSN: 0095-1137  
PUBLISHER: American Society for Microbiology  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB In order to identify antigens that may be used in the serodiagnosis of active tuberculosis (TB), we screened a Mycobacterium tuberculosis genomic expression library with a pool of sera from patients **diagnosed** with active pulmonary TB. The sera used lacked reactivity with a recombinant form of the M. tuberculosis 38-kDa antigen (r38kDa), and the goal was to identify antigens that might complement r38kDa in a serodiagnostic assay. Utilizing this strategy, we identified a gene, previously designated lhp, which encodes a 100-amino-acid protein referred to as culture filtrate protein 10 (CFP-10). The lhp gene is located directly upstream of esat-6, within a region missing in M. bovis BCG. Immunoblot anal. demonstrated that CFP-10 is present in M. tuberculosis CFP, indicating that it is likely a secreted or shed antigen. Purified recombinant CFP-10 (rCFP-10) was shown to be capable of detecting specific antibody in a percentage of TB patients that lack reactivity with r38kDa, most notably in smear-neg. cases, where sensitivity was increased from 21% for r38kDa alone to 40% with the inclusion of rCFP-10. In smear-pos. patient sera, sensitivity was increased from 49% for r38kDa alone to 58% with the inclusion of rCFP-10. In addn., rCFP-10 was shown to be a potent T-cell antigen, eliciting proliferative responses and gamma interferon prodn. from peripheral blood mononuclear cells in 70% of purified protein deriv.-pos. individuals without evident disease. The responses to this antigen argue for the inclusion of rCFP-10 in a polyvalent serodiagnostic test for detection of active TB **infection**. RCFP-10 could also contribute to the development of a recombinant T-cell **diagnostic** test capable of detecting exposure to M. tuberculosis.

REFERENCE COUNT: 25

REFERENCE(S): (1) Alderson, M; J Exp Med 2000, V191, P551 HCAPLUS  
(2) Amara, R; Infect Immun 1996, V64, P3765 HCAPLUS  
(3) Andersen, A; Infect Immun 1989, V57, P2481 HCAPLUS  
(4) Berthet, F; Microbiology 1998, V144, P3195 HCAPLUS  
(6) Chang, Z; J Biol Chem 1994, V269, P1956 HCAPLUS  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 8 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:609008 HCAPLUS

DOCUMENT NUMBER: 133:190202

TITLE: Compounds and methods for the detection and prevention of Trypanosoma cruzi infection

INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Lodes, Michael J.; Houghton, Raymond L.; Smith, John M.; McNeill, Patricia D.

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: PCT Int. Appl., 55 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	----	-----	-----	-----

**THIS PAGE BLANK (USPTO)**

WO 2000050897 A1 20000831 WO 2000-US4815 20000224

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU,  
 CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,  
 IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,  
 MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI,  
 SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM,  
 AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE,  
 DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF,  
 CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.:

US 1999-256976 A 19990224

AB Compds. and methods are provided for **diagnosing** Trypanosoma cruzi **infection**. The disclosed compds. are polypeptides, or antibodies thereto, that contain one or more epitopes of T. cruzi antigens. The compds. are useful in a variety of immunoassays for detecting T. cruzi **infection**. The polypeptide compds. are further useful in vaccines and pharmaceutical compns. for inducing protective immunity against Chagas' disease in individuals exposed to T. cruzi.

REFERENCE COUNT:

7

REFERENCE(S):

- (3) Corixa Corp; WO 9629605 A 1996 HCAPLUS
  - (4) Corixa Corp; WO 9718475 A 1997 HCAPLUS
  - (5) Corixa Corp; WO 9931246 A 1999 HCAPLUS
  - (6) Houghton, R; THE JOURNAL OF INFECTIOUS DISEASES 2000, V181(1), P325 HCAPLUS
  - (7) Reed, S; WO 9316199 A 1993 HCAPLUS
- ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 9 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:588348 HCAPLUS

DOCUMENT NUMBER: 134:290933

TITLE: Cloning of a Mycobacterium tuberculosis gene encoding a purified protein derivative protein that elicits strong tuberculosis-specific delayed-type hypersensitivity

AUTHOR(S):

Coler, Rhea N.; **Skeiky, Yasir A. W.**;  
 Ovendale, Pamela J.; Vedvick, Thomas S.; Gervassi,  
 Luis; Guderian, Jeff; Jen, Shyian; Reed, Steven G.;  
 Campos-Neto, Antonio

CORPORATE SOURCE:

Infectious Disease Research Institute, Seattle, WA,  
 98104, USA

SOURCE:

J. Infect. Dis. (2000), 182(1), 224-233

CODEN: JIDIAQ; ISSN: 0022-1899

PUBLISHER:

University of Chicago Press

DOCUMENT TYPE:

Journal

LANGUAGE:

English

AB The purified protein deriv. (PPD) skin test has been used for the **diagnosis** of tuberculosis for more than 75 yr. However, the test lacks specificity because all mycobacteria share antigens present in PPD. Therefore, sensitization with nontuberculous pathogenic or with environmental nonpathogenic mycobacteria can lead to pos. skin tests. This communication describes a novel PPD protein present only in tuberculous complex mycobacteria. A recombinant protein was obtained and named DPPD on the basis of the first 4 amino acids of its N-terminus sequence. DPPD elicited delayed-type hypersensitivity (DTH) in 100% of Mycobacterium tuberculosis-infected guinea pigs but in no animals sensitized with several organisms representative of all members of the Mycobacterium genus. Preliminary results indicate that DPPD induces strong and specific DTH in humans. This work points to the definition of

**THIS PAGE BLANK (USPTO)**



a single recombinant M. tuberculosis protein that may be an alternative to the PPD test.

REFERENCE COUNT: 44  
 REFERENCE(S): (1) Affronti, L; J Biol Stand 1986, V14, P177 HCAPLUS  
 (2) Affronti, L; J Biol Stand 1986, V14, P187 HCAPLUS  
 (7) Coler, R; J Immunol 1998, V161, P2356 HCAPLUS  
 (9) Daniel, T; Microbiol Rev 1978, V42, P84 HCAPLUS  
 (11) Feinberg, A; Anal Biochem 1983, V132, P6 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 10 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 2000:271863 HCAPLUS  
 DOCUMENT NUMBER: 132:307235  
 TITLE: Combination polypeptide compounds for the detection of Trypanosoma cruzi infection  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Lodes, Michael J.; Houghton, Raymond L.  
 PATENT ASSIGNEE(S): Corixa, USA  
 SOURCE: U.S., 50 pp., Cont.-in-part of Appl. No. PCT/US96/18624.  
 CODEN: USXXAM  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 4  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6054135	A	20000425	US 1997-834306	19970415
WO 9718475	A1	19970522	WO 1996-US18624	19961114
W: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
US 6228372	B1	20010508	US 1997-993674	19971218
PRIORITY APPLN. INFO.:			WO 1996-US18624	A2 19961114
			US 1995-557309	A 19951114
			US 1997-834306	A2 19970415

AB Compds. are provided for **diagnosing** Trypanosoma cruzi **infection**. The disclosed compds. are combination polypeptides that contain a peptide corresponding to novel gene with a repeat sequence and one or more epitopes of T. cruzi antigens such as TcD, TcE, or PEP-2 epitope. The compds. are useful in a variety of immunoassays for detecting T. cruzi **infection**. Peptide epitopes of Trypanosoma cruzi have been identified through expression cloning. A tripeptide (2/D/E) contg. three epitopes (TcD, TcE, PEP-2) was used in ELISA to detect antibodies to T. cruzi in 239 of 240 consensus-pos. sera and 41 of 42 sera confirmed pos. by radioimmunopptn. assay. The 1 discrepant consensus-pos. serum was used to expression-clone a novel T. cruzi gene that contained a repeat sequence. A peptide corresponding to one reading frame of this sequence, TcLol.2, was reactive with T. cruzi-pos. sera and enhanced reactivity of other peptides. Thus, peptides and recombinants contg. multiple repeat epitopes are powerful tools for developing assays for T. cruzi antibody detection and have direct application in blood screening.

**THIS PAGE BLANK (USPTO)**

REFERENCE COUNT: 16  
 REFERENCE(S): (3) Anon; WO 9209895 1992 HCAPLUS  
 (4) Anon; WO 9316199 1993 HCAPLUS  
 (5) Anon; WO 9401776 1994 HCAPLUS  
 (6) Anon; WO 9629605 1996 HCAPLUS  
 (7) Campetella, O; Parasitology Today V8(11), P378  
 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 11 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1999:659510 HCAPLUS  
 DOCUMENT NUMBER: 131:296204  
 TITLE: Fusion proteins of Mycobacterium tuberculosis antigens  
 containing domains from more than one Mycobacterium  
 protein and their uses  
 INVENTOR(S): **Skeiky, Yasir A. W.**; Alderson, Mark;  
 Campos-Neto, Antonio  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 83 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9951748	A2	19991014	WO 1999-US7717	19990407
WO 9951748	A3	20000203		
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
AU 9934817	A1	19991025	AU 1999-34817	19990407
EP 1068329	A2	20010117	EP 1999-916513	19990407
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
BR 9909472	A	20010911	BR 1999-9472	19990407
NO 2000005050	A	20001130	NO 2000-5050	20001006
PRIORITY APPLN. INFO.:			US 1998-56556	A 19980407
			US 1998-223040	A 19981230
			WO 1999-US7717	W 19990407

AB Fusion proteins contg. antigenic regions from two or more proteins (up to five) of Mycobacterium tuberculosis that can be used in the **diagnosis**, treatment and prevention of tuberculosis **infection** are described. These fusion proteins retain the antigenicity of the originals. A series of twelve fusion proteins contg. combinations of peptides from M. tuberculosis antigens were constructed by std. methods and manufd. as inclusion bodies in Escherichia coli. The fusion proteins stimulated T cell proliferation in PPD+ patients with proliferation patterns similar to those of the individual components. Immunization of mice with the fusion proteins induced strong interferon .gamma. and interleukin 4 responses with the strength of the responses depending upon the adjuvant used.

**THIS PAGE BLANK (USPTO)**

L12 ANSWER 12 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1999:549166 HCAPLUS  
 DOCUMENT NUMBER: 131:166234  
 TITLE: Antigenic proteins and their genes for use in  
 diagnosis of tuberculosis  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**;  
 Dillon, Davin C.; Campos-Neto, Antonio; Houghton,  
 Raymond; Vedvick, Thomas S.; Twardzik, Daniel R.;  
 Lodes, Michael J.; Hendrickson, Ronald C.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 323 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9942118	A2	19990826	WO 1999-US3265	19990217
W:	AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
AU 9926819	A1	19990906	AU 1999-26819	19990217
EP 1091749	A1	20010418	EP 1999-907065	19990217
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
PRIORITY APPLN. INFO.:			US 1998-24753 A 19980218	
			US 1998-72596 A 19980505	
			WO 1999-US3265 W 19990217	

AB Numerous antigenic protein fragments are isolated from Mycobacterium tuberculosis culture filtrates and patient sera and used to isolate genomic fragments of the M. tuberculosis DNA. These substances provide reagents for the **diagnosis** of **infection** by detecting antibodies against M. tuberculosis, **diagnosis** of tuberculosis using PCR primers or hybridization probes, and serodiagnosis or immunoassay of tuberculosis. The construction and use of tuberculosis fusion proteins for serodiagnosis are also provided.

L12 ANSWER 13 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1999:549126 HCAPLUS  
 DOCUMENT NUMBER: 131:183862  
 TITLE: Compounds and methods for immunotherapy and diagnosis  
 of tuberculosis  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**;  
 Dillon, Davin C.; Campos-Neto, Antonio; Houghton,  
 Raymond; Vedvick, Thomas S.; Twardzik, Daniel R.;  
 Lodes, Michael J.; Hendrickson, Ronald C.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 299 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English

**THIS PAGE BLANK (USPTO)**

FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9942076	A2	19990826	WO 1999-US3268	19990217
WO 9942076	A3	19991014		
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 9927663	A1	19990906	AU 1999-27663	19990217
EP 1071451	A2	20010131	EP 1999-908169	19990217
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				

PRIORITY APPLN. INFO.:  
US 1998-25197 A 19980218  
US 1998-72967 A 19980505  
WO 1999-US3268 W 19990217

AB Compsds. and methods for inducing protective immunity against tuberculosis are disclosed. The compds. provided include polypeptides that contain at least one immunogenic portion of one or more Mycobacterium tuberculosis proteins and DNA mols. encoding such polypeptides. Such compds. may be formulated into vaccines and/or pharmaceutical compns. for immunization against M. tuberculosis **infection**, or may be used for the **diagnosis** of tuberculosis.

L12 ANSWER 14 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
ACCESSION NUMBER: 1999:537911 HCAPLUS  
DOCUMENT NUMBER: 131:169281  
TITLE: Compounds and methods for the detection of Trypanosoma cruzi infection  
INVENTOR(S): Reed, Steven G.; Houghton, Raymond; **Skeiky, Yasir A. W.**  
PATENT ASSIGNEE(S): Corixa Corporation, USA  
SOURCE: U.S., 26 pp., Cont.-in-part of U. S. 5,756,662.  
CODEN: USXXAM  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5942403	A	19990824	US 1997-929414	19970915
US 5756662	A	19980526	US 1995-403379	19950314
CA 2215104	AA	19960926	CA 1996-2215104	19960312

PRIORITY APPLN. INFO.: US 1995-403379 19950314

AB Compsds. and methods for **diagnosing** Trypanosoma cruzi **infection**, or for screening for T. cruzi or Leishmania **infection**, are disclosed. The disclosed compds. are polypeptides, or antibodies thereto, that contain one or more antigenic epitopes of T. cruzi proteins. The compds. are useful in a variety of immunoassays for detecting T. cruzi **infection**. The polypeptide compds. are further useful in vaccines and pharmaceutical compns. for preventing Chagas' disease in individuals exposed to T. cruzi.

**THIS PAGE BLANK (USPTO)**



REFERENCE COUNT: 20  
 REFERENCE(S): (2) Anon; WO 92/09895 1992 HCAPLUS  
 (3) Anon; WO 93/16199 1993 HCAPLUS  
 (4) Anon; WO 94/01776 1994 HCAPLUS  
 (5) Burns; Proc Nat Acad Sci USA 1992, V89, P1239  
 HCAPLUS  
 (6) Buschiazzi; Molecular and Biochemical Parasitology  
 1992, V54, P125 HCAPLUS  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 15 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1999:405092 HCAPLUS  
 DOCUMENT NUMBER: 131:54723  
 TITLE: Peptides and antibodies for detection and prevention  
 of Trypanosoma cruzi infection  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Lodes,  
 Michael J.; Houghton, Raymond L.; Smith, John M.;  
 McNeill, Patricia D.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 103 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 4  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9931246	A1	19990624	WO 1998-US25871	19981204
W:	AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
US 6228372	B1	20010508	US 1997-993674	19971218
AU 9918057	A1	19990705	AU 1999-18057	19981204
EP 1038000	A1	20000927	EP 1998-962923	19981204
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			

PRIORITY APPLN. INFO.: US 1997-993674 A 19971218  
 US 1997-834306 A2 19970415  
 WO 1998-US25871 W 19981204

AB Polypeptides, or antibodies thereto, that contain one or more epitopes of Trypanosoma cruzi antigens are provided for **diagnosing** T. cruzi **infection**. The compds. are useful in a variety of immunoassays for detecting T. cruzi **infection**. The polypeptide compds. are further useful in vaccines and pharmaceutical compns. for inducing protective immunity against Chagas' disease in individuals exposed to T. cruzi.

REFERENCE COUNT: 3  
 REFERENCE(S): (1) Corixa Corp; WO 9629605 A 1996 HCAPLUS  
 (2) Corixa Corp; WO 9718475 A 1997 HCAPLUS  
 (3) Reed, S; WO 9316199 A 1993 HCAPLUS

L12 ANSWER 16 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1999:350405 HCAPLUS

**THIS PAGE BLANK (USPTO)**

DOCUMENT NUMBER: 131:156679  
TITLE: Molecular characterization and human T-cell responses to a member of a novel Mycobacterium tuberculosis mtb39 gene family  
AUTHOR(S): Dillon, Davin C.; Alderson, Mark R.; Day, Craig H.; Lewinsohn, David M.; Coler, Rhea; Bement, Teresa; Campos-Neto, Antonio; **Skeiky, Y. A. W.**; Orme, Ian M.; Roberts, Alan; Steen, Sean; Dalemans, Wilfried; Badaro, Roberto; Reed, Steven G.  
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA  
SOURCE: Infect. Immun. (1999), 67(6), 2941-2950  
CODEN: INFIBR; ISSN: 0019-9567  
PUBLISHER: American Society for Microbiology  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB The authors have used expression screening of a genomic Mycobacterium tuberculosis library with tuberculosis (TB) patient sera to identify novel genes that may be used **diagnostically** or in the development of a TB vaccine. Using this strategy, the authors have cloned a novel gene, termed mtb39a, that encodes a 39-kDa protein. Mol. characterization revealed that mtb39a is a member of a family of three highly related genes that are conserved among strains of M. tuberculosis and Mycobacterium bovis BCG but not in other mycobacterial species tested. Immunoblot anal. demonstrated the presence of Mtb39A in M. tuberculosis lysate but not in culture filtrate proteins (CFP), indicating that it is not a secreted antigen. This conclusion is strengthened by the observation that a human T-cell clone specific for purified recombinant Mtb39A protein recognized autologous dendritic cells **infected** with TB or pulsed with purified protein deriv. (PPD) but did not respond to M. tuberculosis CFP. Purified recombinant Mtb39A elicited strong T-cell proliferative and gamma interferon responses in peripheral blood mononuclear cells from 9 of 12 PPD-pos. individuals tested, and overlapping peptides were used to identify a min. of 10 distinct T-cell epitopes. Addnl., mice immunized with mtb39a DNA have shown increased protection from M. tuberculosis challenge, as indicated by a redn. of **bacterial** load. The human T-cell responses and initial animal studies provide support for further evaluation of this antigen as a possible component of a subunit vaccine for M. tuberculosis.

REFERENCE COUNT: 44

REFERENCE(S): (1) Abou Zeid, C; Infect Immun 1988, V56, P3046 HCAPLUS  
(2) Amara, R; Infect Immun 1996, V64, P3765 HCAPLUS  
(3) Andersen, P; Infect Immun 1994, V62, P2536 HCAPLUS  
(4) Andersen, P; J Immunol 1995, V154, P3359 HCAPLUS  
(8) Barnes, P; J Immunol 1992, V148, P1835 HCAPLUS  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 17 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1998:790680 HCAPLUS  
DOCUMENT NUMBER: 130:33962  
TITLE: antigenic peptide and nucleic acid sequences useful for diagnosis of tuberculosis and methods for their use  
INVENTOR(S): Alderson, Mark R.; Dillon, Davin C.; **Skeiky, Yasir A. W.**; Campos-Neto, Antonio  
PATENT ASSIGNEE(S): Corixa Corp., USA  
SOURCE: PCT Int. Appl., 104 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent

**THIS PAGE BLANK (USPTO)**

LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 2  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9853076	A2	19981126	WO 1998-US10514	19980520
WO 9853076	A3	19990401		

W: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

US 2001012888	A1	20010809	US 1998-73009	19980505
AU 9875916	A1	19981211	AU 1998-75916	19980520
BR 9809443	A	20000613	BR 1998-9443	19980520
NO 9905690	A	20000118	NO 1999-5690	19991119

PRIORITY APPLN. INFO.: US 1997-858998 A 19970520  
 US 1998-73009 A 19980505  
 WO 1998-US10514 W 19980520

AB Claimed are DNA and protein sequences derived from Mycobacterium tuberculosis antigens, which may be used for **diagnosis** of tuberculosis, and methods for such **diagnosis** with immunol. or genetic methods. The compds. provided include polypeptides that contain at least one antigenic portion of one or more M. tuberculosis proteins, and DNA sequences encoding such polypeptides. **Diagnostic** kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of M. tuberculosis **infection** in patients and biol. samples. Antibodies directed against such polypeptides are also provided.

L12 ANSWER 18 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1998:790679 HCAPLUS  
 DOCUMENT NUMBER: 130:35367  
 TITLE: Mycobacterium antigenic peptides for immunotherapy and diagnosis of tuberculosis and methods of their use  
 INVENTOR(S): Alderson, Mark R.; Dillon, Davin C.; **Skeiky, Yasir A. W.**; Campos-Neto, Antonio  
 PATENT ASSIGNEE(S): Corixa Corp., USA  
 SOURCE: PCT Int. Appl., 100 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9853075	A2	19981126	WO 1998-US10407	19980520
WO 9853075	A3	19990401		

W: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES,

**THIS PAGE BLANK (USPTO)**

FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI,  
CM, GA, GN, ML, MR, NE, SN, TD, TG

AU 9876907 A1 19981211 AU 1998-76907 19980520  
BR 9809445 A 20000613 BR 1998-9445 19980520  
EP 1012293 A2 20000628 EP 1998-924827 19980520  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
IE, FI  
NO 9905689 A 20000118 NO 1999-5689 19991119  
PRIORITY APPLN. INFO.: US 1997-859381 A 19970520  
US 1998-73010 A 19980505  
WO 1998-US10407 W 19980520

AB Mycobacterium antigenic peptides, their nucleotide coding sequences, and their use in vaccines and skin testing for tuberculosis are disclosed. The compds. provided include polypeptides that contain at least one immunogenic portion of one or more M. tuberculosis proteins and DNA mols. encoding such polypeptides. Such compds. may be formulated into vaccines and/or pharmaceutical compns. for immunization against *M. tuberculosis* infection, or may be used for the **diagnosis** of tuberculosis.

L12 ANSWER 19 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
ACCESSION NUMBER: 1998:550516 HCAPLUS  
DOCUMENT NUMBER: 129:187081  
TITLE: Antigens of Leishmania for use in the prophylaxis, diagnosis, and therapy of leishmaniasis  
INVENTOR(S): Reed, Steven G.; Campos-Neto, Antonio; Webb, John R.; Dillon, Davin C.; **Skeiky, Yasir A.**  
PATENT ASSIGNEE(S): Corixa Corporation, USA  
SOURCE: PCT Int. Appl., 195 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9835045	A2	19980813	WO 1998-US3002	19980212
WO 9835045	A3	19981029		
W: BR, MX RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
EP 981624	A2	20000301	EP 1998-907489	19980212
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
BR 9807332	A	20001031	BR 1998-7332	19980212
EP 1113073	A2	20010704	EP 2001-101398	19980212
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
PRIORITY APPLN. INFO.:		US 1997-798841 A 19970212		
		US 1997-920609 A 19970827		
		EP 1998-907489 A3 19980212		
		WO 1998-US3002 W 19980212		

AB Antigens proteins that can be used in the prevention, treatment and **diagnosis** of leishmaniasis are disclosed. CDNAs for antigens of Leishmania major amastigotes were cloned by screening an expression library with antiserum from **infected** mice. The corresponding antigen was manufd. with a hexahistidine affinity label in Escherichia coli and used to immunize mice. Antisera from these mice reacted with all stages of the parasite lifecycle tested (amastigotes, promastigotes).

**THIS PAGE BLANK (USPTO)**



Antigenic peptides of *L. donovani* were identified by their binding to class II MHC antigens and the gene cloned by PCR using amino acid sequence-derived primers. The protein was manufd. as a fusion protein with glutathione-S-transferase and used to immunize rabbits. Antibody from these rabbits recognized a single protein of 23 kilodaltons that appeared to be presented on the cell surface. The protein also stimulated proliferation of T-cells from mice inoculated with *L. donovani* promastigotes. Use of these proteins to induce protective immunity in a mouse exptl. leishmaniasis model.

L12 ANSWER 20 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1998:251273 HCAPLUS

DOCUMENT NUMBER: 128:307522

TITLE: Antigenic polypeptides of *Mycobacterium tuberculosis* and their encoding DNA sequences for immunotherapy and immunotherapy and diagnosis of tuberculosis

INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**;  
Dillon, Davin C.; Campos-Neto, Antonio; Houghton,  
Raymond; Vedvick, Thomas S.; Twardzik, Daniel R.;  
Lodes, Michael J.

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: PCT Int. Appl., 230 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 3

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9816646	A2	19980423	WO 1997-US18293	19971007
WO 9816646	A3	19981008		
W:	AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG			
US 6290969	B1	20010918	US 1997-818112	19970313
AU 9748144	A1	19980511	AU 1997-48144	19971007
EP 932681	A2	19990804	EP 1997-910873	19971007
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
CN 1241212	A	20000112	CN 1997-180501	19971007
BR 9712518	A	20001024	BR 1997-12518	19971007
JP 2001501832	T2	20010213	JP 1998-518456	19971007
NO 9901694	A	19990610	NO 1999-1694	19990409
PRIORITY APPLN. INFO.:			US 1996-730510	A 19961011
			US 1997-818112	A 19970313
			US 1995-523436	B2 19950901
			US 1995-533634	B2 19950922
			US 1996-620874	B2 19960322
			US 1996-659683	A2 19960605
			US 1996-680574	A2 19960712
			WO 1997-US18293	W 19971007

AB Antigenic polypeptides that are useful for the immunotherapy and **diagnosis** of tuberculosis are isolated from the culture filtrate

**THIS PAGE BLANK (USPTO)**

of *M. tuberculosis* strain H37Ra or H37Rv. The antigen-encoding sequences were also isolated from the genomic DNA libraries of the 2 strains. Such compds. may be formulated into vaccines and/or pharmaceutical compns. for immunization against *M. tuberculosis* **infection**, or may be used for the **diagnosis** of tuberculosis. **Diagnostic** kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of *M. tuberculosis* **infection** in patients and biol. samples. Antibodies directed against such polypeptides are also provided.

L12 ANSWER 21 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1998:251272 HCAPLUS

DOCUMENT NUMBER: 128:293966

TITLE: Antigenic polypeptides of *Mycobacterium tuberculosis* and their encoding DNA sequences for diagnosis of tuberculosis

INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**;  
Dillon, Davin C.; Campos-Neto, Antonio; Houghton,  
Raymond; Vedvick, Thomas S.; Twardzik, Daniel R.;  
Lodes, Michael J.

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: PCT Int. Appl., 251 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9816645	A2	19980423	WO 1997-US18214	19971007
W:	AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG			
AU 9747505	A1	19980511	AU 1997-47505	19971007
EP 934415	A2	19990811	EP 1997-910030	19971007
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
CN 1242047	A	20000119	CN 1997-180551	19971007
BR 9712298	A	20001024	BR 1997-12298	19971007
JP 2001500383	T2	20010116	JP 1998-518432	19971007
NO 9901693	A	19990609	NO 1999-1693	19990409
PRIORITY APPLN. INFO.:			US 1996-729622	A 19961011
			US 1997-818111	A 19970313
			WO 1997-US18214	W 19971007

AB Antigenic polypeptides that are useful for the **diagnosis** of tuberculosis are isolated from the culture filtrate of *M. tuberculosis* strain H37Ra or H37Rv. The antigen-encoding sequences were also isolated from the genomic DNA libraries of the 2 strains. **Diagnostic** kits contg. such polypeptides or DNA sequences and a suitable detection reagent may be used for the detection of *M. tuberculosis* **infection** in patients and biol. samples. Antibodies directed against such polypeptides are also provided.

**THIS PAGE BLANK (USPTO)**

L12 ANSWER 22 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1997:443326 HCAPLUS  
 DOCUMENT NUMBER: 127:64504  
 TITLE: Polypeptide epitopes of TcD, TcE or PEP-2 for  
 detection and prevention of Trypanosoma cruzi  
 infection  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Lodes,  
 Michael J.; Houghton, Raymond L.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 110 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 4  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9718475	A1	19970522	WO 1996-US18624	19961114
W:	AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG			
US 5916572	A	19990629	US 1995-557309	19951114
AU 9710568	A1	19970605	AU 1997-10568	19961114
AU 722152	B2	20000720		
EP 874992	A1	19981104	EP 1996-941416	19961114
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
JP 2000503111	T2	20000314	JP 1997-519161	19961114
BR 9611455	A	20010102	BR 1996-11455	19961114
US 6054135	A	20000425	US 1997-834306	19970415
PRIORITY APPLN. INFO.:			US 1995-557309 A	19951114
			WO 1996-US18624 W	19961114

AB Compds. and methods are provided for **diagnosing** Trypanosoma cruzi **infection**. The disclosed compds. are polypeptides contg. TcD, TcE or PEP-2 antigen epitope of Trypanosoma cruzi, or antibodies to the antigens. The polypeptides are coupled to nitrocellulose, latex, or plastic material, and used with labeled binding agent (e.g. anti-Ig., protein G, protein A or lectins) for immunoassay of Trypanosoma cruzi in blood, serum, plasma, saliva, cerebrospinal fluid or urine samples. The polypeptide compds. are further useful in vaccines and pharmaceutical compns. for inducing protective immunity against Chagas' disease in individuals exposed to T. cruzi.

L12 ANSWER 23 OF 27 HCAPLUS COPYRIGHT 2001 ACS  
 ACCESSION NUMBER: 1997:309979 HCAPLUS  
 DOCUMENT NUMBER: 126:276349  
 TITLE: Immunogenic Mycobacterium tuberculosis antigens and genes for immunotherapy and diagnosis of tuberculosis  
 INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**;  
 Dillon, Davin C.; Campos-Neto, Antonio; Houghton,  
 Raymond; Vedvick, Thomas H.; Twardzik, Daniel R.  
 PATENT ASSIGNEE(S): Corixa Corporation, USA  
 SOURCE: PCT Int. Appl., 167 pp.

**THIS PAGE BLANK (USPTO)**

CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 3  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9709428	A2	19970313	WO 1996-US14674	19960830
WO 9709428	A3	19970717		
W: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML				
CA 2230885	AA	19970313	CA 1996-2230885	19960830
ZA 9607394	A	19970505	ZA 1996-7394	19960830
EP 851927	A2	19980708	EP 1996-933009	19960830
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, FI				
CN 1200147	A	19981125	CN 1996-197639	19960830
BR 9610262	A	19990706	BR 1996-10262	19960830
JP 2001517069	T2	20011002	JP 1997-511464	19960830
AU 9671586	A1	19970327	AU 1996-71586	19960930
AU 727602	B2	20001214		
NO 9800883	A	19980427	NO 1998-883	19980227

## PRIORITY APPLN. INFO.:

US 1995-523436	A	19950901
US 1995-533634	A	19950922
US 1996-620874	A	19960322
US 1996-659683	A	19960605
US 1996-680574	A	19960712
WO 1996-US14674	W	19960830

AB Compds. and methods for inducing protective immunity against tuberculosis are disclosed. The compds. provided include polypeptides that contain at least one immunogenic portion of one or more M. tuberculosis antigens and DNA mols. encoding such proteins. Such compds. may be formulated into vaccines and/or pharmaceutical compns. for immunization against M. tuberculosis **infection**, or may be used for the **diagnosis** of tuberculosis.

L12 ANSWER 24 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1997:296911 HCAPLUS

DOCUMENT NUMBER: 126:276348

TITLE: Antigenic polypeptides of Mycobacterium tuberculosis and their encoding DNA sequences for diagnosis of tuberculosis

INVENTOR(S): Reed, Steven G.; **Skeiky, Yasir A. W.**; Dillon, Davin C.; Campos-Neto, Antonio; Houghton, Raymond; Vedvick, Thomas H.; Twardzik, Daniel R.

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: PCT Int. Appl., 189 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

**THIS PAGE BLANK (USPTO)**



PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9709429	A2	19970313	WO 1996-US14675	19960830
WO 9709429	A3	19970717		
W: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML				
CA 2230927	AA	19970313	CA 1996-2230927	19960830
AU 9671587	A1	19970327	AU 1996-71587	19960830
ZA 9607395	A	19970424	ZA 1996-7395	19960830
EP 850305	A2	19980701	EP 1996-933010	19960830
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
CN 1200146	A	19981125	CN 1996-197467	19960830
BR 9610268	A	19990706	BR 1996-10268	19960830
JP 11514217	T2	19991207	JP 1996-511465	19960830
PRIORITY APPLN. INFO.:				
			US 1995-523435	A 19950901
			US 1995-532136	A 19950922
			US 1996-620280	A 19960322
			US 1996-658800	A 19960605
			US 1996-680573	A 19960712
			WO 1996-US14675	W 19960830
AB Antigenic polypeptides that useful for the diagnosis of tuberculosis are isolated from the culture filtrate of M. tuberculosis strain HH37Ra or H37Rv. The antigen-encoding sequences were also isolated from the genomic DNA libraries of the 2 strains. Diagnostic kits contg. such polypeptides or DNA sequences and a suitable detection reagent for the detection of M.tuberculosis in biol. samples are provided. Antibodies directed against such polypeptides are also provided.				
L12 ANSWER 25 OF 27 HCAPLUS COPYRIGHT 2001 ACS				
ACCESSION NUMBER: 1995:891909 HCAPLUS				
DOCUMENT NUMBER: 124:46736				
TITLE: Cloning and expression of a Leishmania donovani gene instructed by a peptide isolated from major histocompatibility complex class II molecules of infected macrophages				
AUTHOR(S): Campos-Neto, Antonio; Soong, Lynn; Cordova, Jose L.; Sant'Angelo, Derek; <b>Skeiky, Yasir A. W.</b> ; Ruddle, Nancy H.; Reed, Steven G.; Janeway, Charles, Jr.; McMahon-Pratt, Diane				
CORPORATE SOURCE: Medical Sch. Ribeirao Preto, Univ. Sao Paulo, Sao Paulo, Brazil				
SOURCE: J. Exp. Med. (1995), 182(5), 1423-33				
CODEN: JEMEAV; ISSN: 0022-1007				
DOCUMENT TYPE: Journal				
LANGUAGE: English				
AB The studies reported here describe the isolation of peptides from MHC class II mols. of murine macrophages <b>infected</b> with Leishmania donovani, and the use of the derived peptide sequences to rescue the pathogen peptide donor protein. The isolation of the peptides was carried out by comparing the RP HPLC profile of peptides extd. from <b>infected</b> macrophages with the peptides extd. from noninfected cells. Several distinct HPLC peaks unique to <b>infected</b> macrophages was sequenced. One of the peptides that was not homologous to				

**THIS PAGE BLANK (USPTO)**

any known protein was used to instruct the designing of an oligonucleotide sense primer that was used in combination with an oligo dT nucleotide (anti-sense primer) to amplify by PCR a DNA fragment from *L. donovani* cDNA. The amplified DNA fragment was cloned and used as a probe to screen a *L. donovani* cDNA library. The cloned gene (Ld peptide gene) has an open reading frame of 525 bp and has no homol. with any known protein/gene sequence. Northern blot analyses indicated that the Ld peptide/gene is broadly distributed and expressed among species of the *Leishmania* genus, in both the amastigote and promastigote life cycle forms. Using the pGEX 2T vector, the gene was expressed and the relationship of the purified recombinant protein with *L. donovani* was confirmed using both antibody and T cell responses from immunized or **infected** animals. The gene encodes a 23-kD mol. (Ldp 23) assocd. with the cell surface of *L. donovani* promastigotes. In addn., T cells purified from the lymph nodes of BALB/c mice immunized with *L. donovani* or **infected** with *L. major*, and from CBA/J mice **infected** with *L. amazonensis* were stimulated to proliferate by the recombinant Ldp 23 and produced high levels of IFN- $\gamma$  and no IL 4. This observation suggests that the Ldp 23 is an interesting parasite mol. for the studies concerning the host-parasite interaction because the Th1 pattern of cytokine response that it induces is correlated with resistance to *Leishmania* **infections**. These results clearly point to an alternative strategy for the purifn. of proteins useful for the development of both vaccines and immunol. **diagnostic** tools not only against leishmaniasis but also for other diseases caused by intracellular pathogens.

L12 ANSWER 26 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1991:580954 HCAPLUS

DOCUMENT NUMBER: 115:180954

TITLE: Immunological characterization of recombinant antigens isolated from a *Mycobacterium avium* .lambda.gt11 expression library by using monoclonal antibody probes  
AUTHOR(S): Rouse, David A.; Morris, Sheldon L.; Karpas, Arthur B.; Mackall, Julia C.; **Probst, Peter G.**;  
Chaparas, Sotiros D.

CORPORATE SOURCE: Cent. Biol. Eval. Res., Food Drug Adm., Bethesda, MD, 20892, USA

SOURCE: Infect. Immun. (1991), 59(8), 2595-600

CODEN: INFIBR; ISSN: 0019-9567

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Nontuberculous mycobacteria, particularly *M. avium*, have been isolated from a significant percentage of patients with AIDS. Early detection of *M. avium* **infection** is difficult, and treatment regimens are often ineffective. Much needs to be learned about antigens and factors responsible for immunity to and pathogenesis of the disease. Specific antigens and **diagnostic** procedures for **infection** need to be developed. To address some of these problems, monoclonal antibodies were generated against a serovar 4 strain of *M. avium* isolated from a patient with AIDS. Protease sensitivity studies have demonstrated that each of these antibodies recognizes a protein-assocd. epitope. Immunoblot analyses suggest that seven of these monoclonal antibodies react specifically with *M. avium* and *M. intracellulare* epitopes. Immunoreactive **bacteriophages** were identified from an *M. avium* .lambda.gt11 expression library with two of these monoclonal antibodies (3808 C3 and 3954 B12). Lambda lysogens, generated from the immunoreactive **bacteriophages**, overproduced .beta.-galactosidase fusion proteins which were reactive with the two monoclonal antibodies in immunoblot assays. The purified fusion proteins were shown to elicit skin test

**THIS PAGE BLANK (USPTO)**

reactions in sensitized guinea pigs.

L12 ANSWER 27 OF 27 HCAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1990:457041 HCAPLUS

DOCUMENT NUMBER: 113:57041

TITLE: Production, characterization, and species specificity of monoclonal antibodies to Mycobacterium avium complex protein antigens

AUTHOR(S): Rouse, David A.; Morris, Sheldon L.; Karpas, Arthur B.; Probst, Peter G.; Chaparas, S. D.

CORPORATE SOURCE: Cent. Biol. Eval. Res., Food Drug Adm., Bethesda, MD, 20892, USA

SOURCE: Infect. Immun. (1990), 58(5), 1445-9

CODEN: INFIBR; ISSN: 0019-9567

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The incidence of M. avium-M. intracellulare complex **infections** has increased in recent years primarily because a significant proportion of acquired immunodeficiency syndrome patients develop disseminated M. avium complex disease. In an effort to develop new tools to study these **infections**, 8 monoclonal antibodies directed against M. avium were produced. Western blot (immunoblot) specificity anal. and protease sensitivity assays indicate that 4 of these antibodies recognize M. avium-specific protein epitopes and 2 react with M. avium complex-specific peptide determinants. These monoclonal antibodies may be useful clin. in the **diagnosis** of M. avium complex disease and in the lab. for isolation and characterization of native and recombinant M. avium complex antigens.

**THIS PAGE BLANK (USPTO)**